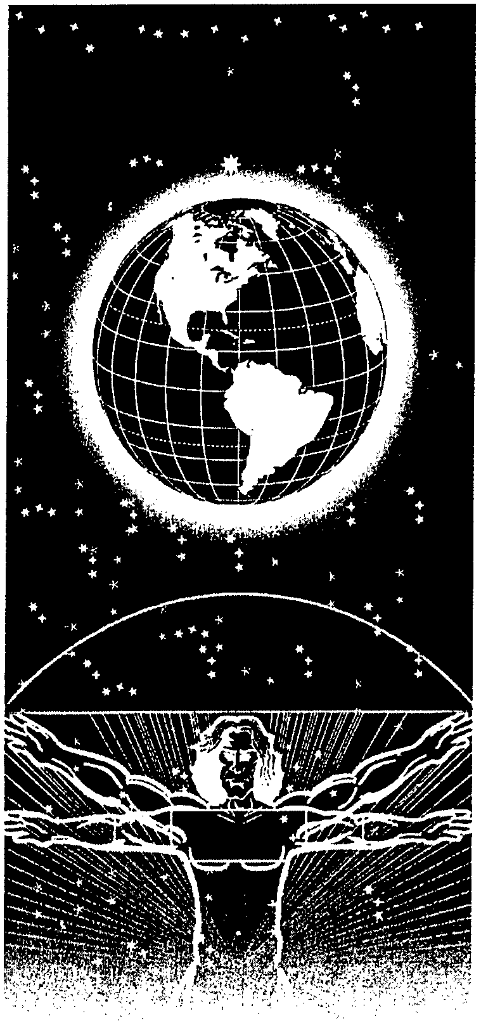
AFRL-HE-BR-TR-2003-0086



UNITED STATES AIR FORCE RESEARCH LABORATORY

THE EFFECTS OF 37 HOURS OF CONTINUOUS WAKEFULNESS ON THE PHYSIOLOGICAL AROUSAL, COGNITIVE PERFORMANCE, SELF-REPORTED MOOD, AND SIMULATOR FLIGHT PERFROMANCE OF F-117A PILOTS

John Caldwell
J. Lynn Caldwell
David Brown
Nicholas Smythe
Jennifer Smith
Jeffery Mylar
Monica Mandichak
Christina Schroeder

HUMAN EFFECTIVENESS DIRECTORATE
BIODYNAMICS AND PROTECTION DIVISION
FLIGHT MOTION BRANCH
2485 GILLINGHAM DRIVE
Brooks City-Base, TX 78235

June 2003

NOTICES

This report is published in the interest of scientific and technical information exchange and does not constitute approval or disapproval of its ideas or findings.

This report is published as received and has not been edited by the publication staff of the Air Force Research Laboratory.

Using Government drawings, specifications, or other data included in this document for any purpose other than Government-related procurement does not in any way obligate the US Government. The fact that the Government formulated or supplied the drawings, specifications, or other data, does not license the holder or any other person or corporation, or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

The Office of Public Affairs has reviewed this paper, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This report has been reviewed and is approved for publication.

John Caldwell

JOHN A. CALDWELL, Ph.D.
Project Scientist

F. Wesley B

F. WESLEY BAUMGARDNER, Ph.D.
Chief, Biodynamics and Protection Division

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
June 2003

2. REPORT TYPE
Interim

3. DATES COVERED *(From - To)*
01-03-2003- -16-06-2003

4. TITLE AND SUBTITLE
The Effects of 37 Hours of Continuous Wakefulness on the Physiological Arousal, Cognitive Performance, Self- Reported Mood, and Simulator Flight Performance of F-117A pilots

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
62202F

6. AUTHOR(S)
John Caldwell, J. Lynn Caldwell, David Brown, Nicholas Smythe, Jennifer Smith, Jeffery Mylar, Monica Mandichak and Christina Schroeder

5d. PROJECT NUMBER
7757

5e. TASK NUMBER
P9

5f. WORK UNIT NUMBER
04

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Reseach Laboratorv
Human Effectiveness Directorate
Biodynamics and Protection Division
Flight Motion Branch
2485 Gillingham Dr.
Brooks City-Base, TX 78235

8. PERFORMING ORGANIZATION REPORT NUMBER
AFRL-HE-BR-TR-2003-0086

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/HE

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release, distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Over the past 30 years, fatigue has contributed to a number of Air Force mishaps. Resource cutbacks combined with increased operational tempos, sustained operations, and night fighting could exacerbate the problem. Extended wakefulness and circadian factors can be especially problematic in military aviation where mission demands sometimes necessitate flights as long as 17-44 hours. To effectively counter fatigue in such operations, the effects of this threat must be objectively measured and understood. This study assessed F-117A pilots during a 37-hour period of continuous wakefulness. Although none of the pilots crashed, substantial decrements were found in flight-skills, reaction times, mood, and brain activation as early as after the 26th hour of continuous wakefulness. The greatest flight degradations occurred after 27-33 hours awake, even though many pilots believed their worst performance was earlier. The decrements found in this relatively-benign test environment may be more serious under operational conditions unless personnel anticipate the most dangerous times and administer valid fatigue countermeasures.

15. SUBJECT TERMS
Fatigue, Sustained Operations, Fighter Pilots, Aviation

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
					John Caldwell
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**	Unclass	81	**19b. TELEPHONE NUMBER** *(include area code)*
Unclass	Unclass	Unclass			210-536-6371

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

BACKGROUND

Over the past 20 years, the contribution of fatigue to air incidents and accidents has increasingly been recognized in both the civil and military aviation communities. As the following section will illustrate, numerous studies have documented the general problems associated with impaired alertness in flight, and several investigations have identified the primary underlying causes of overly tired pilots. Although little work has focused specifically on the direct impact of fatigue on basic piloting skill in modern Air Force aircraft, it is clear that the fatigue-related degradations that theoretically form the underpinnings of pilot performance are increasingly being acknowledged for their negative impact on operational readiness. Along with this recognition, there has been an increased focus on designing tools or techniques capable of reliably measuring, monitoring, and predicting fatigue levels in the aviation and other militarily-important environments. The ultimate aim of much of this work is to identify personnel who are at risk for fatigue-related failures early on so that they either can be removed from harm's way or provided with scientifically-valid fatigue countermeasures before their impaired alertness poses a hazard to mission safety or effectiveness. A clear understanding of the full spectrum of fatigue effects in conjunction with timely alertness predictions and remedies for degraded abilities is key to ensuring optimal pilot performance across the broad spectrum of military aviation requirements.

The Reasons for Concern over Pilot Fatigue

Pilot fatigue is an insidious threat, especially in operations involving sleep loss from circadian disruptions and extended duty periods, as well as those in which personnel must work during the nighttime or predawn hours when the body's natural rhythms are known to impair arousal (Akerstedt, 1995). Needless to say, these factors are all common in aviation

sustained operations where "24/7" schedules are often essential for effective mission completion. Technological advances such as night vision devices have enhanced the night-fighting capabilities of both ground and air troops, making around-the-clock operations a highly feasible component of the modern military strategy. Combining efficient day and night fighting capabilities across successive 24-hour periods places a significant strain on enemy resources and presents a clear tactical advantage for U.S. forces. In fact, the Air Force Chief of Staff recently noted that persistent and sustained operations "24 hours a day, seven days a week…" are essential to attaining U.S. victory in today's battlespace (Elliot, 2001). In addition, "owning the night" offers the advantage of utilizing the cover of darkness to capitalize on the element of surprise as well as to maintain the secrecy of tactical and other mission-related strategies or activities.

Unfortunately, however, there are difficulties inherent in maintaining effective around-the-clock operations. Aircraft and other equipment can function nonstop for extended periods without adverse effects, but human operators need periodic sleep for the restoration of both the body and the brain (Horne, 1978). Depriving humans of proper restorative sleep rapidly produces high levels of fatigue that can lead to dangerously impaired performance (Krueger, 1989). Aviator fatigue is associated with degradations in response accuracy and speed, the unconscious acceptance of lower standards of performance, impairments in the capacity to integrate information, and narrowing of attention that can lead to forgetting or ignoring important aspects of flight tasks (Perry, 1974). Fatigued pilots tend to decrease their physical activity, withdraw from social interactions, and lose the ability to effectively divide mental resources among different tasks. As sleepiness levels increase, performance becomes less consistent and overall vigilance deteriorates (Dinges, 1990). It has been determined that sleep-deprived personnel lose approximately 25 percent of their higher-level mental

capabilities with each 24-hour period of sleep loss (Belenky et al., 1994). Furthermore, scientific studies have established that as little as 17-24 hours of sustained wakefulness can produce psychomotor deficits equivalent to those observed with blood alcohol concentrations (BAC's) of 0.05% to 0.10% (Dawson and Reid, 1997).

Pilots flying at night or during the predawn hours are especially vulnerable to fatigue-related cognitive lapses, or even worse, "micro-sleeps"—brief periods during which sleep uncontrollably and often unconsciously intrudes into wakefulness. Moore-Ede (1993) found that aviators engaged in simulator flights during the predawn hours experienced a tenfold increase in the number of inadvertent sleep lapses, and during this same time (in which the micro-sleeps were most prominent), pilots made the greatest number of performance errors. Klein, Bruner, and Holtman (1970) reported that pilots' abilities to fly a simulator at night decreased to a level comparable to that observed with a BAC of 0.05%. Wright and McGown (2001) found that while sleepiness (both during the daytime and the night) was increasingly problematic as a function of increased flight duration, occurrences of outright sleep episodes were most frequent on night flights with late-night departures compared to flights that departed earlier in the day. These findings corroborated those of Rosekind et al. (1994) who had previously concluded that pilots were particularly plagued by inadvertent (and often unrecognized) periods of dozing off as well as concurrent decrements in vigilance performance and subjective alertness ratings during night flights in comparison to flights during the daytime hours.

In light of these findings, it is easy to understand why pilot fatigue has been identified as a substantial threat to flight safety. In commercial civilian aviation, overly-tired pilots are thought to be responsible for 4-7 percent of air incidents or accidents every year in the U.S. (Kirsh, 1996). NASA's Aviation Safety Reporting System (ASRS) routinely receives reports

from pilots blaming fatigue, sleep loss, and sleepiness in the cockpit for operational errors such as altitude and course deviations, fuel miscalculations, landings without proper clearances, and landings on incorrect runways (Rosekind et al., 1994). In the military realm, pilot fatigue is no less problematic. A recent report identified fatigue as a contributing factor in 4 percent of Army aviation mishaps from 1990-1999 (personal communication, U.S. Army Safety Center help desk; helpdesk@safetycenter.army.mil). In the Air Force, aircrew fatigue has been at least partially blamed for 7.8 percent of the reportable Class A mishaps that have occurred over the past 30 years (personal communication, Lt Col Thomas Luna, U.S. Air Force Safety Center). In addition, 25% of the Air Force's night tactical fighter Class A accidents were attributed to fatigue between 1974 and 1992, and 12.2% of the Navy's total Class A mishaps were thought to be the result of aircrew fatigue from 1977 to 1990 (Ramsey and McGlohn,1997).

Summary of the Role of Fatigue

Taken together, these statistics make it clear that the fatigue from mission-related sleep deprivation and circadian factors must be completely understood and appreciated for the dangers that it may pose in operational contexts. Not only must the *general* characteristics of aviator fatigue be elucidated, as has been done within the context of the previously-cited investigations, but the *specific* effects of fatigue on basic pilot flight-performance skills and other aspects of pilot functional status must be further explored by assessing the actual flight performance of current and qualified military aviators flying specially-instrumented weapon system trainers and specially-instrumented aircraft. Such data are crucial to the development and implementation of scientifically valid fatigue countermeasures that can be explicitly tailored to the demands of modern military fixed-wing operations. In addition, it is essential to explore effective strategies for reliably measuring the degree to which pilot fatigue is

present on a moment-to-moment basis so that the proper timing of fatigue countermeasures can be optimized across a wide range of militarily-relevant circumstances.

Interest in Predicting Fatigue-Related Impairments

There has been a long-standing interest in developing the capability to make accurate predictions regarding the cognitive and performance status of human operators in the workplace. The U.S. Army Medical Research and Materiel Command has for several years focused a large-scale effort towards accomplishing this goal in its Warfighter Physiological Status Monitoring (WPSM) Program, and the National Academies' Institute of Medicine is currently conducting a comprehensive review of whether or not it is possible to monitor the performance status of military personnel in operational contexts. The reason behind such efforts is that, obviously, if it were possible to reliably monitor and evaluate the real-time cognitive and physical status of personnel, it would be possible to continuously maximize the overall performance of entire units by identifying "weak links" and resolving identified problems in a timely fashion. Overly-tired troops could be removed from the battle before they made potentially disastrous mistakes. Sleepy pilots could be warned about impending alertness decrements in time to administer effective fatigue countermeasures. Busy commanders could optimize their overall mission performance with the aid of computerized scheduling assistants that included physiological status data. In short, military personnel could monitor and plan around their physiological limitations the same way that automobile and truck drivers now plan their next rest and recuperation stops based on the oil-pressure readings and fuel gauges of their vehicles. From the standpoint of mental readiness, some type of on-line alertness monitor would prevent personnel from unexpectedly "running out of gas" at the times when optimal performance is needed most.

In the search for a feasible way to accomplish such status monitoring or status predictions in pilots, several factors must be considered. First and foremost, a decision must be made about whether the goal is to make a preflight prediction of upcoming performance or whether the objective is to make continuous in-flight assessments about what may happen in the next few seconds.

Preflight Predictions of Upcoming Status—A Role for Eye Movement Measures

If the goal is to predict a pilot's suitability for an upcoming flight, eye-movement measures may offer important information. The sensitivity of the oculomotor control system to fatigue, boredom, and lapses in attention has previously been established by several investigators. Long duration eye closures during eyeblinks have been found to provide one indication of reduced alertness (Stern, 1980), and recent studies have demonstrated that an integrated measure of the degree of eye closure over a specific interval of time offers information about the level of operator sleepiness (Wierwille, 1999). In fact, Dinges et al. (1998) found a high degree of coherence between slow eye closures and performance lapses on a test of sustained attention. Stern (1999) has shown that several other oculomotor measures also may be useful for detecting lapses in attention. Among these are the saccadic eye movements that quickly transition the eyes from one point of focus to another. The latency between stimulus presentation and the saccade as well as the duration of the saccade have been suggested as indicators of fatigue. Although saccades are initiated under voluntary control, once they are underway, their speed is not within the individual's control. Fatigued conditions can cause a longer latency period, reduce the velocity of the saccade, or result in saccades that either undershoot or overshoot the target. Russo et al. (2003) found that saccadic velocity is particularly sensitive to an increase in sleepiness in response to prolonged periods of partial sleep deprivation.

Pupil diameter is another eye measure that has shown promise for assessing general levels of fatigue. According to Stern (1999), a decrease in pupil diameter and/or a slow fluctuation in pupil diameter can coincide with feelings of tiredness. Similarly, Russo et al. (1999) found that both decreases in saccadic velocity and increases in pupil constriction latency correlated with an increase in the rate of crashes under simulated driving conditions during periods of sleep deprivation. Russo et al. (2003) later summarized information which suggested that the speed and magnitude with which the pupil contracts after exposure to a light flash may offer insight into the level of parasympathetic/sympathetic-nervous-system dominance in fatigued volunteers. Thus, there is reason to believe that a variety of eye-movement measures may offer reliable predictions about the upcoming performance capabilities of pilots. Of course, such measures cannot be evaluated in flight because the already heavy demands on the visual system would make standardized in-flight testing impossible. However, pre-flight testing appears to be a viable option. For this reason, devices such as the Fitness Impairment Tester (PMI, Inc., 1999) have been developed to measure saccadic velocity, pupil diameter, pupil-contraction latency, and pupil-constriction amplitude in an effort to reliably assess changes in operator status attributable to fatigue, stress, and drugs.

Continuous, Real-Time Status Assessments—A Role for EEG Measures

If continuous, real-time assessment is the goal, the chosen measure of pilot status must be based upon data that are attainable during the completion of the operational flight tasks. In addition, the collection/analysis methodology must be realistically feasible considering the current state of equipment technology, and the selected measure must be objective, valid, and demonstrably related to important changes in operational readiness. For these reasons, it appears that the electroencephalogram (EEG) is most promising for pilot status monitoring

(Caldwell et al., 1994b; 1997; 2001). The EEG can be collected without interfering with the primary task, and it is the most direct indication of the CNS functioning that is the foundation of higher-order mental activity. Numerous studies have established the sensitivity of EEG activity to work-related stressors such as sleep deprivation (Comperatore et al., 1993 ; Caldwell et al., 1996, Caldwell et al., 2003 ; Lorenzo et al., 1995 ; Pigeau et al., 1987; and others). In general, these investigations have shown that slow-wave EEG activity (i.e., delta and/or theta) is significantly elevated by even moderate (18-20 hours) sleep loss, and such changes would be expected to correlate with performance decrements since it is well known that this level of sleep deprivation is associated with a variety of mental impairments. Another important point about the EEG is that it has already been successfully collected from aviators in the in-flight environment. LaFontaine and Medvedeff (1966), Maulsby (1966), and Howitt et al. (1978) have offered evidence of the utility of using the EEG as a general status measure during flights, and Sterman et al. (1987) indicated that increased flying demands were associated with elevations in frontal EEG theta power, reductions in frontal EEG alpha power, and increased EEG asymmetries between left and right central regions. Caldwell et al. (1994a; 1997; 2001) have shown that it is feasible to collect and telemeter spontaneous EEG even from the high-vibration environment found in helicopter flights. In addition, Caldwell et al. (2001) demonstrated that standard fatigue-related changes in EEG data associated with sleep loss (elevations in theta power and decreases in alpha power) could be recorded, under resting conditions (with a safety pilot on the controls), from pilots flying in a UH-60 helicopter. Unfortunately, no studies are currently available that have substantiated a reliable predictive relationship between in-flight EEG changes and actual fluctuations in pilot performance. However, it is reasonable to postulate that such a relationship does in fact exist since research has repeatedly established that the fatigue from

insufficient sleep affects EEG activity and non-concurrently-measured flight performance (Caldwell et al., 2001; LeDuc, Caldwell, and Ruyak, 2000; Caldwell et al., 2000a; Caldwell et al., 2000b).

Summary of Fatigue Assessment Strategies

The available data clearly establish a potential role both for pre-flight fatigue assessments based on oculomotor measures as well in-flight continuous assessments based on electroencephalographic measures of CNS activity. If it can be established that these status-assessment strategies are reliable indicators of actual pilot performance, general pilot alertness, and basic aviator cognition, they can be fielded to objectively optimize operational work/rest schedules as well as to guide the timely implementation of validated fatigue countermeasures in aviation sustained operations.

RATIONALE FOR THE PRESENT STUDY

A host of evidence suggests that aircrew fatigue poses significant risks in high-performance military flight operations. However, at the present time, there are no published studies that precisely detail the impact of pilot fatigue on the objectively measured flight performance of pilots flying any of the U.S. Air Force's modern fighter or bomber aircraft such as the F-15, the F-117, the B-1, or the B-2. Furthermore, there are no studies that have objectively examined the impact of fatigue on flight performance in the weapon system trainers for these aircraft. The present study will partially fill this void by evaluating the impact of sustained wakefulness and night flying on the flight performance, cognition, and alertness of F-117A pilots. This effort is necessary to provide baseline measures of the impact of untreated fatigue so that, in the near future, the operational utility of both non-pharmacological and pharmacological fatigue remedies can be determined.

In addition, the current study will explore potential methods for predicting impending decrements in pilot performance in a way that will optimize the accurate and timely implementation of fatigue countermeasures. At present, there are apparently no published studies that have assessed fatigue-related changes in the central nervous system (CNS) activation of Air Force pilots concurrently performing flight maneuvers in a weapons training simulator of any of the actual aircraft that make up the current equipment inventory. In addition, there are no studies that have empirically established whether a specific set of oculomotor measures of CNS status (those contained in the Fitness Impairment Tester™ mentioned earlier) may be useful for predicting the onset of fatigue-related flight-performance decrements, although it has been assumed by some that this test will be useful in Air Force flight operations.

OBJECTIVES

The principal objective of this research is to establish the effects of long duty periods, sleep deprivation, and night missions on the flight performance, mood, cognition, and physiological activation of a subset of U.S. Air Force pilots. Accomplishment of this objective will provide insight into the exact nature and timing of the fatigue-related decrements that will likely occur in sustained fighter operations. Specifically, this investigation will establish the impact of 37 hours of sustained wakefulness on:

- *objectively-measured pilot performance* in F-117A pilots flying standardized flight maneuvers in a specially-instrumented flight simulator,
- *central nervous system (CNS) arousal* based on electroencephalographic (EEG) assessments of the amounts of delta, theta, and alpha activity,
- *parasympathetic/sympathetic activation* based upon measures of pupil diameter, constriction amplitude, constriction latency, and saccadic velocity,

- *self-reported measures of psychological mood states*, alertness, sleepiness, energy, and other aspects of subjective status,
- and *general cognitive status* in terms of the ability to perform simple mathematical evaluations as well as aviation-related divided-attention tasks.

In addition, this research will determine the degree to which changes in objectively-measured simulator flight performance coincide with changes in EEG and oculomotor parameters. Accomplishment of this later objective will provide an indication of whether these specific assessments of aviator physiological status hold promise for guiding the implementation of operational fatigue countermeasures.

METHODS

Subjects

Ten qualified pilots (mean age of 35.7 years, ranging from 27-43 years old) who were members of the 49th Fighter Wing at Holloman Air Force Base, NM, served as participants after signing an informed consent agreement which detailed the procedures of the study. All participants were in possession of a current "up slip" (Air Force Form 1042, medical permission to engage in military flying duties) at the time of their admission, and all were current and qualified in the F-117A aircraft. The participants possessed an average of 2,431 total flight hours (ranging from 929 to 4900 hours) and 426 F-117 flight hours (ranging from 42 to 1035 hours). Five of the volunteers were dual-rated in a T-38 aircraft. All of the participants were male because there currently are no female F-117A pilots. Only those who had been working the day shift (those normally remaining awake from approximately 0600 or 0700 to 2100 or 2300) for at least 5 days prior to the study were allowed to enroll. None were taking any type of medication known to impact mental alertness (sedating antihistamines, sleep medications, prescription stimulants, etc.). No restrictions on

experience were imposed because this factor has not been shown to impact the resistance to sleep deprivation, nor has it been shown to affect the relationship between performance capacity and electrophysiological variables in studies conducted over the past 16 years by the principal investigator of the present research. A total of 10 pilots were evaluated because this number of participants was determined to yield sufficient statistical power based on power analyses conducted on data from an earlier study which was performed on helicopter pilots at the U.S. Army Aeromedical Research Laboratory, Fort Rucker, AL.

Apparatus

The entire research protocol was conducted inside of the F-117 Weapon System Training (WST) facility at Holloman Air Force Base, NM. The flight-performance data were collected with the simulator and ancillary equipment. The remaining measures were collected with various laboratory testing devices which were set up in a co-located sound-attenuated testing room (within the simulator facility).

Multi-Attribute Test Battery (MATB)

The MATB (Comstock and Arnegard, 1992) is a computerized aviation simulation test that required participants to perform an unstable tracking task while concurrently monitoring warning lights and dials, responding to computer-generated auditory requests to adjust radio frequencies, and managing simulated fuel flow rates (using various key presses). This test was controlled by a Micron Pentium-based computer equipped with a standard keyboard, a joystick, and a mouse. Data on tracking errors, response times, time-outs, false alarms, and accuracy rates were calculated via the use of the MATB processing software.

Mathematical Processing

The Mathematical Processing subtest from the Automated Neuropsychological Assessment Metrics (ANAM) battery (Reeves et al., 1993) is a basic cognitive test that

required participants to solve arithmetic problems that were presented in the middle of the computer screen. The task involved deducing an answer to an equation such as "5 + 3 - 4 =" and then deciding if the answer was greater-than or less-than the number five. Based on the calculation, the participant then pressed one of two specified response buttons (on the mouse). This test was controlled by a standard Pentium-based desktop computer equipped with a keyboard and a mouse (which was used to make the required responses to each item). Data on performance accuracy, response speed, and throughput were calculated by computer via STATVIEW™ software at the conclusion of testing.

FIT Workplace Safety Screening Evaluation

The Fitness Impairment Tester (PMI, Inc., 1999) is a computerized fitness-for-duty test that required participants to peer into a device in which visual stimuli (both moving and stationary) were presented. The device detects changes in pupil size (as small as 0.05 mm) and movements of the eye (as small as one degree) in response to controlled flashes of light and moving light targets. Measures of saccadic velocity, pupil diameter, pupil-contraction latency, and pupil-constriction amplitude were calculated by the FIT device and then downloaded to a Pentium-based computer for processing in a relational database.

EEG/Electro-oculographic (EOG)/Electrocardiographic (EKG) Recordings

Physiological data were collected with a Grass-Telefactor Instruments Aurora recording system (West Warwick, RI) running TWin™ collection and analysis software. For the EEG data, Grass gold-cup electrodes filled with Mevidon electrolyte gel were used (21 EEG channels were referenced to A1 and A2 during recording). For the EOG data, Grass F-E9M-60-5 11-mm Silver/Silver Chloride electrodes filled with Grass EC2 electrolyte paste were used. For the EKG data, Kendall MediTrace disposable, self-adhesive EKG electrodes were used. Data were digitized at a rate of 200 samples per second. The recording filters were set at 1.0-70 Hz for the EEG, 0.3-35 Hz for the EOG, and 1.0-35 Hz for the EKG. During all data collection (whether in the simulator or in the co-located testing room), the quality of the recordings were monitored continuously in real time in an effort to make corrections of any problems which were encountered (i.e., excessive body/eye movements or muscle artifact). Nevertheless, in the case of the data recorded in the simulator, all activity above the alpha band (more than 13 Hz) ultimately was disregarded from analysis due to the presence of muscle tension that could not be eliminated while the pilots were actively concentrating on the flight tasks.

Profile of Mood States

Subjective evaluations of mood were made with the Profile of Mood States (POMS) (McNair, Lorr, and Droppleman, 1981). The POMS is a 65-item questionnaire which, when scored according to the specified templates, measures affect or mood on 6 scales: 1) tension-anxiety, 2) depression-dejection, 3) anger-hostility, 4) vigor-activity, 5) fatigue-inertia, and 6) confusion-bewilderment. Factor scores on each scale are analyzed.

Visual Analog Scales

In addition to the POMS, subjective sleepiness and alertness (and other parameters) were measured via the Visual Analog Scale (VAS) (an adaptation of the one developed by Penetar et al., 1993). This questionnaire consists of several 100-millimeter lines, each of which is labeled at one end with the words "not at all" and at the other end with the word "extremely." Centered under each line are the test adjectives which are as follows: “alert/able to concentrate,” “anxious,” “energetic,” “feel confident,” “irritable,” “jittery/nervous,” “sleepy,” and “talkative.” The participant indicates the point on the line which corresponded to how he felt along the specified continuum at the time at which the test is taken. The score for each item consisted of the number of millimeters from the left side of the line to the location at which the participant placed his mark.

Flight Simulator

The F-117A Weapon System Trainer (L-3 Communications/Link Simulation and Training, Binghamton, NY) was used to conduct all of the flight-performance assessments. The Weapon System Trainer (WST) is a stationary digital device that simulates the characteristics and operations of the F-117A stealth fighter aircraft that is currently in the U.S. Air Force equipment inventory. The WST provides a fully-functioning replica of the interior cockpit of the actual aircraft, including all primary and secondary flight controls, aural cues (engine sounds), and cockpit lighting (L-3 Communications, 1993). The components of the WST include the simulator itself as well as an instructor/operator station (IOS), a computer complex that includes an Alpha Server 8200 and Input/Ouput (I/O) cabinets, and the equipment necessary for the generation of out-of-the-window and IR visual scenes. The actual F-117A aircraft (simulated by this WST) is a twin-turbofan powered, low-radar, ground-attack fighter with a single-seat cockpit. The F-117A WST faithfully simulates the F-117A aircraft to the extent that training in

the WST is directly transferable in terms of instrument flights, takeoffs and landings, instrument navigation, system operations, and air-to-ground attack procedures. In the present study, only the instrument-flight simulation capability was utilized.

All WST flights were set up for night illumination conditions with zero visibility and no visible lighting on the horizon. This was done to ensure that all pilots remained focused on the flight instruments (simulated Instrument-Flight-Rules conditions) throughout the entire test period. In addition, the WST was set up to generate zero air turbulence with no wind gusts in order to prevent non-pilot-related flight-path deviations. The auto-throttle and auto-pilot modes were disengaged to force all participants to "hand fly" the simulator.

Objective flight performance data were collected using the Coherent Automated Simulation Test Environment (CoASTE) tool—a set of software routines that normally provide the capability to evaluate simulator performance, display/manipulate various data from simulator data pools, and/or trace and correct problems. The CoASTE's trace utility was used to capture various parameters of flight performance data (see Table 1) at a rate of 2 Hz throughout each flight. One complete data file was generated for each simulator flight, and this file contained all of the data collected from the beginning to the end of the given simulation session. Each record in the file contained the time at which each data sample was collected, the actual data points themselves, and an identification field which consisted of the subject number, the testing day, and the testing session. The completed data files were downloaded to a Read/Write Compact Disk (CD) at the conclusion of data collection before being transferred to a standard desktop Pentium-based computer where each file was segmented into the individual maneuvers that comprised the overall flight profile. Afterwards, root mean square (RMS) errors for maneuver-relevant parameters were calculated for statistical analysis.

The measures (data points) recorded for flight-performance data analysis are shown in the table below. The individual flight maneuvers (and the measures scored for each) are later described in the Procedures section of this report.

Table 1. Measured Simulator Flight Parameters

Parameter	Range
1. Indicated altitude	0-30,000 feet
2. Indicated airspeed	30-600 KIAS
3. Indicated vertical speed	0 +/- 5,000 fpm
4. Magnetic heading	0-360 degrees
5. Pitch angle	0 +/- 90 degrees
6. Roll angle	0 +/- 90 degrees
7. Slip	0 +/- 2 balls
8. Localizer/course deviation	0 +/- 2 dots
9. Glideslope/course deviation	0 +/- 2 dots

WAM (Wrist Activity Monitors)

Wrist monitors (Ambulatory Monitoring, Inc., Ardsley, NY) were used to track sleep/activity rhythms in a relatively unobtrusive fashion. In this study, the WAMs (which are battery-powered devices about the size of a wrist watch) were used only to motivate subjects to follow admonishments not to sleep from 0600 on the morning of their test day until the time at which they reported to the Laboratory for testing (i.e., at 1800). Activity data were downloaded once the participant arrived at the simulator facility for electrode application (prior to the sleep-deprivation period), and computer-generated actigraphs were visually inspected to ensure compliance with the "no sleep" rule. The data from the WAMs were not further analyzed.

Procedure

Each participant completed three training/familiarity sessions on the first day of his participation. Then, starting on the following day and continuing through the night and throughout the next day, each participant completed five testing sessions that covered the final 23 hours of a 37-hour period of continuous wakefulness (the participants had already been awake for 14 hours before the first deprivation test began).

Prior to being admitted to the study, each participant's medical records were screened for current illnesses or disqualifying medications by the medical monitor or his designee at the Holloman AFB medical clinic. Afterwards, for training, participants arrived at approximately 1230 to sign the informed consent agreement and to be briefed on all of the upcoming study procedures. One of the Associate Investigators met one-on-one with each participant to address any questions or concerns that the participants may have had. Once the informed consent agreement was signed, the member of the research staff who served as the console operator during all of the initial training flights informed the participant about the flight maneuvers that would be flown and the manner in which each flight would be conducted (i.e., it was explained that a staff member would explicitly sequence the volunteer through each of 15 flight maneuvers at the appropriate point in each of the flight profiles).

Training Schedule

Training flights were conducted at approximately 1400, 1700, and 2000. Although the exact timing of these flights was not considered critical, since their only purpose was to ensure asymptotic (or near-asymptotic) levels of proficiency on the maneuvers to be used during the actual deprivation test flights, the timing of the training flights generally was within 30 minutes of the target flight times. During each of the training flights, objective

flight-performance data were collected, but the majority of these data were disregarded (since the training curves were not of particular interest in this study).

Interspersed among the training flights, there were six iterations of the MATB, nine iterations of the mathematical processing test, three iterations of the POMS/VAS, and two FIT evaluations that were conducted for familiarization/training/baseline purposes. Once all of these tests were completed (by approximately 2100 on the training day), each volunteer received a WAM with instructions on how to place it on his wrist. He was then reminded to wake up at 0600 the next morning and to avoid any napping after 0700, prior to returning for deprivation testing at 1800 the next day. Furthermore, he was cautioned not to ingest any type of caffeinated product between 1000 and his 1800-hour report time. Afterwards, he was released for the evening.

Testing Schedule

On the testing day, participants arrived at the simulator facility at 1800 for electrode application. Twenty-five scalp placements were marked according to the International 10-20 system for electrode placement, each placement site was thoroughly cleaned with acetone, and EEG electrodes were attached to the scalp with collodion. After attachment, each EEG electrode was filled with electrolyte gel (through a small hole in the top of the electrode), and then impedances were checked to ensure they were below 5000 Ohms. In addition, 2 self-adhesive ECG electrodes were attached to the participant's sternum and rib cage (one each), and EOG electrodes were attached above and below the left eye with adhesive collars (EOG electrodes were filled with electrolyte paste prior to attachment to the skin since these are closed electrodes which cannot be filled after placement).

Once all of the electrodes were attached, the participant proceeded to his first EEG test which was conducted at 2100. For this evaluation, the participant was seated in the

designated testing area (which was quiet and isolated) where he was connected to the EEG recording equipment and instructed to sit quietly for 2 minutes with eyes open followed by 2 minutes with eyes closed. Following EEG testing, the participant completed one POMS and one VAS. Next, he performed the MATB for 30 minutes. Afterwards, the participant completed the FIT, mathematical processing test, and another resting EEG, POMS, and VAS.

Once the non-flight testing for the session was complete, the participant was escorted into the simulator at which time he completed an eyes-open/eyes-closed EEG while seated in the simulator (4 minutes total) prior to beginning the flight profile. Afterwards, the participant completed the maneuvers listed in Table 2. The same staff member instructed the participant exactly when to begin each maneuver during each of the deprivation flights; however, the participants were tasked with ending the maneuvers at the correct altitudes and headings. If participants correctly maintained ideal performance parameters, all maneuvers were flown at an airspeed of 300 knots, all climbs and descents were flown at a climb/descent rate of 1000 feet per minute, and all turns were flown at a 30-degree angle of bank. Flight performance and EEG activity (as well as EOG and EKG data) were recorded continuously.

Approximately 2 hours after the conclusion of the flight, the participant began the next non-flight test session (which contained another series of EEG, POMS, VAS, MATB, FIT, EEG, POMS, and VAS evaluations) at 0200. At 0400, the participant started the second flight. After this flight, there were three more non-flight test sessions (at 0700, 1200 and 1700) and three more simulator flights (at 0900, 1400 and 1900). Following the 1900 flight, the participant's electrodes were removed, and he was debriefed and then released.

Table 2. Flight maneuvers

Number	Detailed maneuver descriptions
1.	Right 360° turn at an altitude of 11,000 feet mean sea level (MSL)
2.	Straight and level on a heading of 345 degrees at 11,000 feet MSL
3.	Left 360° turn at an altitude of 11,000 feet MSL
4.	Straight climb from 11,000 to 13,000 feet MSL
5.	Straight and level on a heading of 345 degrees at 13,000 feet MSL
6.	Descending right 360° turn to an altitude of 10,000 feet MSL
7.	Straight and level on a heading of 345 degrees
8.	Left-climbing 540° turn to an altitude of 15,000 feet.
9.	Straight and level on a heading of 165 degrees at 15,000 feet MSL
10.	Right 360° turn at an altitude of 15,000 feet MSL
11.	Straight and level on a heading of 165 degrees at 15,000 feet MSL
12.	Left 720° turn at an altitude of 15,000 feet MSL
13.	Straight descent from 15,000 to 13,000 feet MSL
14.	Intercept localizer (not scored)
15.	Instrument Landing System (ILS) approach to Runway 16

In summary, there were five non-flight sessions and five flights during the last hours of the 37-hour period of continuous wakefulness. The non-flight sessions were conducted at 5-hour intervals beginning at 2100, 0200, 0700, 1200 and 1700. The simulator flights were conducted at 5-hour intervals following the other test sessions at 2300, 0400, 0900, 1400 and 1900. The entire routine is detailed below in table 3.

Table 3. Testing routine across the training day and the deprivation period.

Time of Day	Day 1 Training	Day 2 Testing	Day 3 Testing
0200 0400		 **Wakeup**	EEG/POMS/VAS/MATB ANAM/FIT/EEG/POMS/VAS FLIGHT
0700 0900			EEG/POMS/VAS/MATB ANAM/FIT/EEG/POMS/VAS FLIGHT
1200 1400	*Informed Consent* FLIGHT MATB/ANAM (twice) POMS/VAS/FIT		EEG/POMS/VAS/MATB ANAM/FIT/EEG/POMS/VAS FLIGHT
1700 1900	FLIGHT MATB/ANAM (twice) POMS/VAS FLIGHT MATB/ANAM (twice) POMS/VAS	***Electrode Application***	EEG/POMS/VAS/MATB ANAM/FIT/EEG/POMS/VAS FLIGHT
2100 2300	Don Wrist Monitor Release	EEG/POMS/VAS/MATB ANAM/FIT/EEG/POMS/VAS FLIGHT	***Release***

DATA ANALYSIS

As noted earlier, there were two iterations of EEG, POMS, and VAS testing during each non-flight test session. In this report, only the tests that were performed closest to the simulator flights were analyzed.

Missing data were handled via BMDPAM which substituted the cell means for any missing observations (there was never more than 5 percent of data loss due to equipment failures or any other problems). Afterwards, the data (non-flight and flight) were initially analyzed with BMDP4V, Repeated Measures Analysis of Variance (ANOVA). Huyhn-Feldt adjusted degrees of freedom were used to compensate for any observed violations of the compound symmetry assumption that is critical for accuracy in repeated-measures ANOVA (note that this correction results in fractional degrees-of-freedom values). Follow-up (post-hoc) tests most frequently consisted of regression evaluations for the presence of linear,

quadratic, and cubic trends (also calculated with BMDP4V). Trend analysis was used in place of pairwise contrasts in an effort to minimize alpha inflation (for five test times, only three trend analyses were necessary as opposed to the 24 posthoc tests that would have been necessary for all possible pairwise comparisons). For most of the data, one-way ANOVAs were used since there were five iterations of testing and no other design factors. However, the EEG data and the flight-performance data required analysis with a two-way ANOVAs because there were two design factors of interest in these data sets. For the resting EEGs, data were collected at five testing times both under eyes-open and under eyes-closed conditions. For the simulator EEGs (collected while the participant was flying the maneuvers), the data from two different straight-and-levels were analyzed (straight-and-levels 3 and 5) at each of the deprivation testing times. With the flight-performance data, there were five testing iterations and eight maneuver types (combined right 360° turns, combined straight-and-levels, left 360° turn, straight climb, straight descent, left-climbing turn, left 720° turn, and right descending turn).[1] These maneuver types were analyzed together in a maneuver-by-time ANOVA after they were analyzed separately in a series of one-way ANOVAs. To combine the two right turns with one another, the performance scores were simply averaged prior to analysis after determining that there were no differences between the two iterations of this maneuver. To combine the five straight-and-levels, the same approach was used, again after it was determined that there were no differences between the five iterations of this maneuver.

The flight performance data collected from the 2300, 0400, 0900, 1400, and 1900 flights were initially converted into RMS errors prior to their conversion into percentage-of-change-

[1] The ILS approach will not be analyzed at this point due to technical issues which have not yet been resolved.

from-baseline (i.e., change from the 1900-hour training-day flight, as described below). The RMS errors were based upon different parameters from the various maneuvers because, for instance, it would make little sense to evaluate heading deviations during turns since the heading in a turn necessarily must change in order to accomplish the maneuver. The specific parameters that were evaluated for each of the maneuvers are depicted in table 4. The RMS errors were calculated by subtracting the observed moment-to-moment parameter values (sampled at a rate of 2 Hz) from the target value for that particular parameter.[2] This deviation score was squared, then summed across the entire maneuver. The result was divided by the total number of samples collected during the maneuver, and then, the square root of this result was obtained (the formula for calculating RMS errors is the same as the one used to calculate standard deviations except that deviations from the target value rather than deviations from the mean are of interest).

All data (flight and non-flight, with the exception of the EEG data which were not collected prior to the deprivation period) were converted into scores that represented the "percentage change from baseline" prior to analysis. This was accomplished in a series of steps. First, for each iteration of each test, the score that was derived from the last iteration of this same test on the previous training day (baseline) was subtracted from the score that was derived during each of the deprivation-testing iterations (for the flight data, the score was subtracted from the baseline since the flight data consisted of RMS errors, and we wished to show changes in accuracy rather than increases in errors). Next, the results of these calculations were divided by the baseline score. Finally, the outcomes were each

[2] To minimize the chances of RMS errors becoming inflated due to improper maneuver set up rather than to increased control variability, all participants were required to be within 2 degrees of the target heading, 5 degrees of the target airspeed, and 20 feet of the target altitude before each maneuver was started.

multiplied by 100. Thus, the formula used for all of the non-flight tests (except the EEGs) was as follows:

$$\textbf{\textit{Percent Change = ((Score – Baseline)/Baseline) x 100.00.}}$$

Whereas the percent-change-from-baseline score for the RMS errors stemming from the simulator flight data was:

$$\textbf{\textit{Percent Change = ((Baseline - Score)/Baseline) x 100.00.}}$$

For example, the POMS fatigue score earned by a participant during his 2230 POMS at the outset of the deprivation period was subtracted from the POMS fatigue score he earned earlier on the previous training day at 2100. The result was then divided by the training-day score (collected at 2100 the previous day) and multiplied by 100. The same series of calculations was performed for the POMS fatigue scores earned at each subsequent deprivation session (0330, 0830, 1330, and 1830). In this way, it was possible to calculate the percentage increase (or decrease) in self-rated fatigue throughout the deprivation period. This same basic procedure was used on all of the other data as well (except for the EEG).

The strategy of calculating and analyzing percent-change scores not only provided an intuitively-appealing method for examining the effects of fatigue on all of the individual laboratory tests conducted in this research, it also offered an avenue for calculating composite flight scores for each of the flight maneuvers. This is because normally it would not be possible to combine measures of different flight parameters such as headings, altitudes, and airspeeds due to the fact that each of these are based on different scales (i.e., headings are expressed in degrees, altitudes are expressed in feet, and airspeeds are expressed in knots). However, by converting each of the measured flight parameters to a *percentage of change from baseline*, (using the last training flight as the baseline) all of the individual parameter scores (headings, altitudes, airspeeds, etc.)--now expressed as percentages--could

be averaged together to represent overall maneuver performance during each of the deprivation flights. Had this not been the case, it would have been necessary to individually analyze each relevant parameter from each of the flight maneuvers, and this would have created a problem with alpha-rate inflation due to the large number of statistical comparisons required. In fact, with 3 relevant measures per maneuver type, and with the relevant measures changing as a function of the particular requirements of each maneuver, a minimum of 24 separate ANOVAs would have been necessary as opposed to the 8 ANOVAs which permitted analysis of each maneuver separately plus a single additional ANOVA which enabled the evaluation of the composite percent-change scores from all maneuvers combined.

The method for analyzing the flight data was as follows: once all RMS errors had been calculated and all flights for a particular subject were complete, change-from-baseline scores were calculated for each of three relevant parameters (i.e., heading, altitude, airspeed, etc.) for each maneuver. Finally, the 3 percent-change-scores for each maneuver were averaged together to yield one composite change-from-baseline score for each maneuver within each of the deprivation flights. For the combined maneuver analysis, these data were analyzed in a two-way ANOVA as explained earlier. In addition, each maneuver was analyzed separately to examine the consistency of effects across all eight maneuver types.

Table 4. Parameters evaluated in each of the flight maneuvers.

Maneuver	Altitude	Airspeed	Vertical Velocity	Heading	Roll
Right 360° Turn No. 1	x	x			x
Straight-and-Level No. 1	x	x		x	
Left 360° Turn	x	x			x
Climb		x	x	x	
Straight-and-Level No. 2	x	x		x	
Right Descending Turn		x	x		x
Straight-and-Level No. 3	x	x		x	
Left Climbing Turn		x	x		x
Straight-and-Level No. 4	x	x		x	
Right 360° Turn No. 2	x	x			x
Straight-and-Level No. 5	x	x		x	
Left 720° Turn	x	x			x
Descent		x	x	x	

EEG data were classified into the 4 standard activity bands of delta (1.5-3.0 Hz), theta (3.0-8.0 Hz), alpha (8.0-13.0 Hz), and beta (13.0-20 Hz) by calculating the power spectrum (using a Hamming window) on a minimum of three 2.5-second epochs within each EEG segment of interest. This procedure was used for both the resting EEG collected in the co-located testing room (prior to each flight) and the resting EEG collected in the simulator prior to the initiation of the flight maneuvers. The same was done for the EEG segments that were collected during each subjects' performance of each of the flight maneuvers. Each band was analyzed in a two-way, repeated measures analysis of variance (ANOVA) in which the

factors are as follows: 1) for the laboratory data (collected in the co-located testing room), the factors were time (2200, 0300, 0800, 1300, 1800) and eyes (eyes open, eyes closed); 2) for the in-flight data, the factors were time (2300, 0400, 0900, 1400, and 1900) and section (resting eyes-open in the simulator, EEG collected during SL-3, and EEG collected during SL-5). Note that in the present report, only the EEG data collected immediately prior to each flight has been examined. The remaining EEG data will be the subject of a future report.

The EKG and EOG data collected during the flight testing in this investigation have been stored on disk and transferred to Dr. Glenn Wilson at Wright Patterson Air Force Base where they first will be pre-processed (converted into blinks per minute, beats per minute, and heart-rate variability) and then subjected to two-way ANOVAs for time and maneuver. A report on the outcome of this analysis is forthcoming.

The percent-change scores from each of the items on the VAS (i.e., sleepiness, alertness, energy, etc.) and each of the factor scores on the POMS (i.e., fatigue, confusion, vigor, etc.) were analyzed with a series of one-way ANOVAs across deprivation testing times. Only the self-report/mood data that were collected closest to the flights were examined in this report. Each set of scores from both the VAS and the POMS were analyzed separately.

The percent-change scores on the reaction times, time outs, hits, misses, and tracking errors from the four subtests of the MATB (communications, system monitoring, resource management, and tracking) also were analyzed separately with a series of one-way ANOVAs for time (2130, 0230, 0730, 1230, and 1730).

The percent-change scores from the FIT Workplace Safety Screener (pupil diameter, constriction latency, constriction amplitude, and saccadic velocity) were each subjected to a one-way ANOVA for time (2210, 0310, 0810, 1310, and 1810).

Following the analyses for fatigue effects, bivariate correlations were calculated on a subset of the variables that seemed likely to be sensitive to sleep deprivation. This was done to determine whether EEG, oculomotor, MATB, or self-report data could be used to predict changes in flight performance. Using BMDP2R, the composite flight scores from each session were correlated with corresponding saccadic velocity, resting eyes-open EEG theta activity (collected under laboratory conditions) at Cz, MATB tracking errors, POMS fatigue, ANAM throughput, VAS alertness, and the in-flight EEG data recorded from Cz during the resting, SL-3, and SL-5 sections of the flights (within each individual participant). Percent-change measures were used (with the exception of the EEG). Afterwards, overall correlations were calculated by averaging the correlations obtained for each set of measures across all 10 participants.

RESULTS

Fatigue Effects

The cognitive data, physiological measures, subjective self-report data, and objective flight performance data were evaluated with independent repeated-measures ANOVAs. All significant effects were then examined for the presence of linear, quadratic, and cubic trends.

Multi-Attribute Task Battery

The MATB consists of four concurrently-administered subtests that were evaluated during each test session. The subtests were: communications, systems monitoring, fuel monitoring, and unstable tracking. As described below, a variety of performance indicators were collected on each of the subtests, and prior to analysis, these data were transformed into percent-change-from-baseline scores, using the data from the 2030 test on the training day as the baseline. Following the percent-change transformation, each MATB measure was

analyzed in a separate one-way ANOVA to determine whether there were differences across the various sleep-deprivation testing times (2130, 0230, 0730, 1230, and 1730).

Communications. In the communications task, participants responded to simulated radio calls by dialing in a designated "radio frequency." The reaction time (RT), standard deviation of reaction times (SDRT), accuracy, and time-out (TO) errors were analyzed. Of these measures, none were found to be sensitive to the effects of sleep loss in F-117 pilots.

Systems monitoring. In the monitoring task, participants monitored various warning lights and dials, and their RT for each response, the SDRT, and the TO errors (missed signals) were assessed during each iteration of the task. Out of these three measures, one was found to be sensitive to the effects of 37 hours of continuous wakefulness. There was a significant main effect on the time factor for the overall RT ($F(4,36)=4.58$, $p=.0043$). This was due to a linear slowing in the RT to lights and dials ($p<.05$). See figure 1.

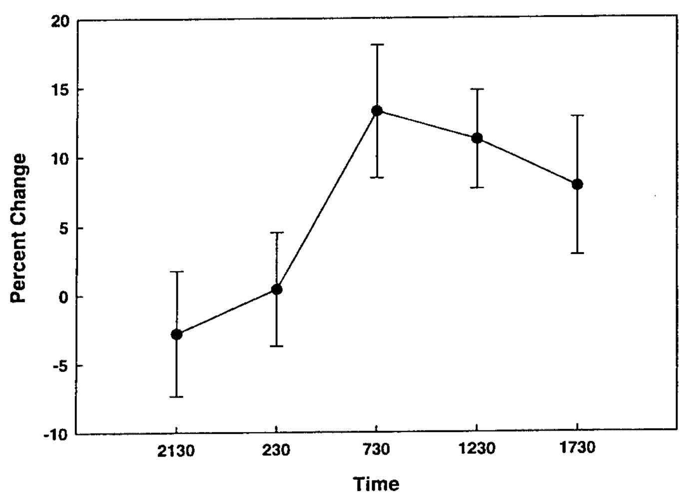


Figure 1. The effects of continuous wakefulness on MATB communications RT (expressed as percentage of change from baseline).

Fuel monitoring. In addition to monitoring lights and dials, participants were required to maintain the fuel levels in simulated fuel tanks at 2500 pounds. The one-way ANOVA on

the absolute deviation from this target value indicated that there were no differences in the ability to accomplish this subtask as a function of sleep deprivation.

Unstable tracking. The tracking task required participants to maintain an unstable target at a point in the middle of the tracking window while performing the other subtasks mentioned above. One-way ANOVA on tracking RMS errors indicated that the participants' ability to accomplish this objective was affected by 37 hours of continuous wakefulness ($F(2.47,22.24)=7.65$, $p=.0018$). As the graphic depiction in figure 2 shows, performance was linearly degraded overall ($p<.05$), and this was augmented by a particularly-noticeable increase in tracking errors from 2130 to 0730 (after which there was a leveling off until the end of the deprivation period).

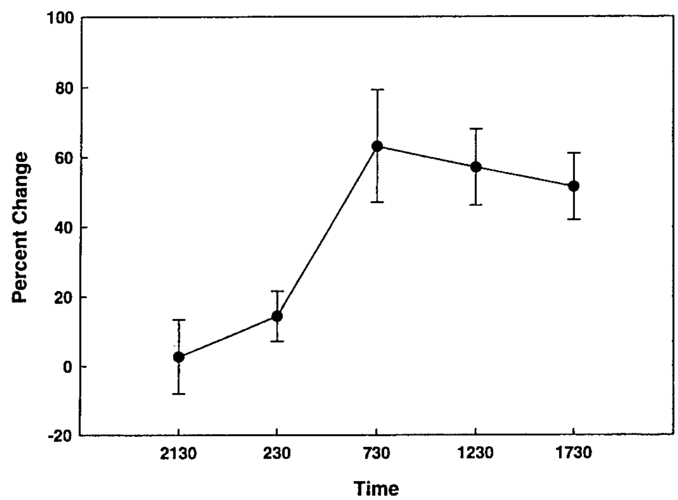


Figure 2. The effects of continuous wakefulness on MATB RMS tracking errors (expressed as percentage of change from baseline).

Mathematical Processing (ANAM)

The mathematical processing task from the ANAM battery was analyzed in terms of RT for correct responses, SDRT for correct responses, overall accuracy, and number of correct responses per minute (throughput). Prior to analysis, these data were transformed into percent-change-from-baseline scores, using the data from the 2055 test on the training day as

the baseline. Following the percent-change transformation, each measure was analyzed in a separate one-way ANOVA to determine whether there were differences across the various sleep-deprivation testing times (2205, 0305, 0805, 1305, and 1805).

Reaction time for correct responses. The one-way ANOVA on the RT data revealed overall differences among the testing times ($F(4,36)=5.16$, $p=.0022$). Trend analysis indicated there was a significant quadratic change ($p<.05$) that occurred because of a sharp slowing of RT from 2205 to 0805 followed by a partial recovery at 1305 and 1805 (see figure 3).

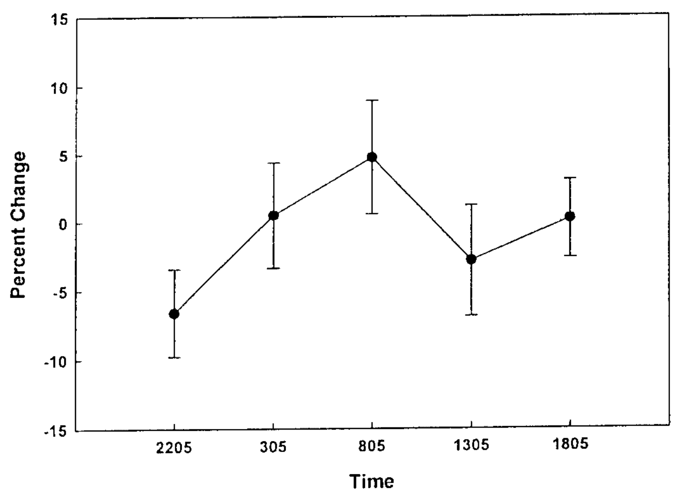


Figure 3. The effects of continuous wakefulness on ANAM mathematical processing reaction time (expressed as percentage of change from baseline).

Standard deviation of RT for correct responses. The analysis of the RT variability data indicated there was not a significant main effect as a function of test time (sleep deprivation). Apparently, this measure was not sensitive to the level of sleep loss to which the volunteers in this study were exposed.

Accuracy. The accuracy data also were unaffected by sleep deprivation as indicated by the absence of statistically-significant differences among the testing sessions.

Throughput. The number of correct responses per minute (a combined speed/accuracy measure) indicated marked changes throughout the deprivation period ($F(4,36)=4.97$, $p=.0027$). Subsequent trend analysis revealed a significant linear decrease from the first to the last test session ($p<.05$) and a cubic pattern of reductions from 2205 to 0805 followed by a throughput recovery from 0805 to 1305 after which there was another decline from 1305 to 1805 ($p<.05$). These effects are graphically depicted in figure 4.

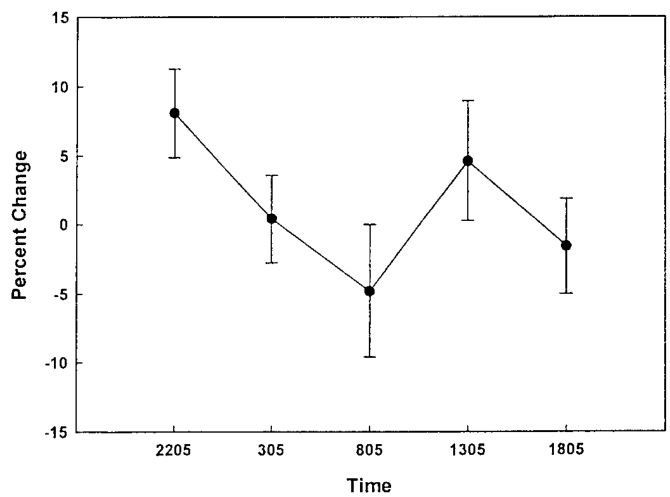


Figure 4. The effects of continuous wakefulness on ANAM mathematical processing throughput (expressed as percentage of change from baseline).

Oculomotor (FIT) data

There were four oculomotor measures collected during each of the five equally-spaced test sessions. The measures were pupil diameter, pupil constriction amplitude, pupil constriction latency, and saccadic velocity. Prior to analysis, these data were transformed into percent-change-from-baseline scores using data from the 1320 test on the training day as the baseline (there were no other FIT tests after this time on the training day). Following the data transformations, each oculomotor measure was analyzed in a separate one-way ANOVA to determine whether there were differences across the various sleep-deprivation testing times (2210, 0310, 0810, 1310, and 1810).

Pupil diameter. Measurements of the participants' initial pupil diameters were unchanged across the various testing times from 2210 to 1810.

Pupil constriction amplitude. The degree to which the participants' pupils constricted in response to brief flashes of light also were unaffected by the level of sleep deprivation experienced in this study.

Pupil constriction latency. The amount of time from the onset of light flashes until pupil constriction tended to differ across the deprivation testing times, although statistical significance was not obtained ($p=.0564$). However, in light of the strength of this tendency, subsequent trend analyses were performed and the data were graphed. The trend analysis revealed a quadratic change ($p<.05$) which was attributable to the fact that constriction latencies grew slower from 2210 to 0310, remained slow from 0310 to approximately 1310, and then returned (at 1810) to approximately the same speed as was observed during the first test at 2210 (see figure 5).

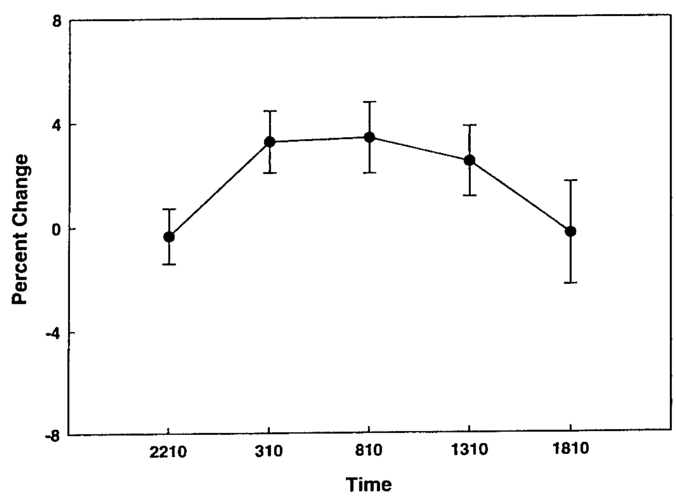


Figure 5. The effects of continuous wakefulness on pupil constriction latency (expressed as percentage of change from baseline).

Saccadic velocity. The speed of saccades were substantially affected by sleep deprivation as indicated by a significant time main effect ($F(2.17,19.50)=5.18$, $p=.0140$.

Trend analysis revealed that there was a quadratic change in the data ($p<.05$). Saccadic velocity decreased substantially from 2210 to 0310, remained relatively slow from 0310 to 1310, and then increased at 1810 (see figure 6).

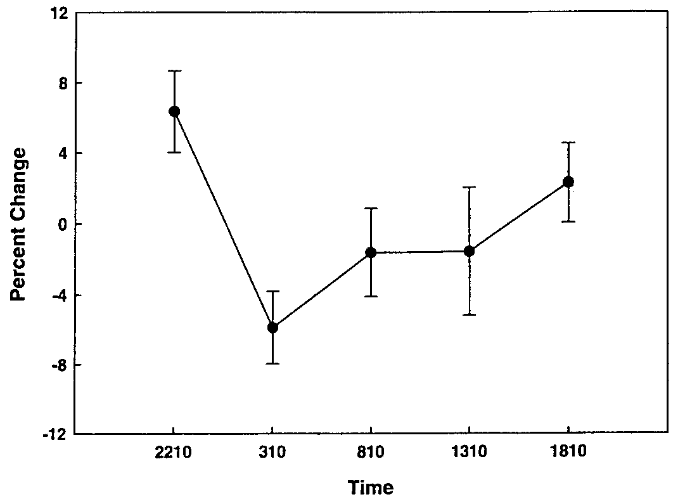


Figure 6. The effects of continuous wakefulness on saccadic velocity(expressed as percentage of change from baseline).

Resting EEG

Resting EEG data were collected under eyes-open and eyes-closed conditions (2 minutes of each) at each of the deprivation sessions. From each of the 2-minute segments, a minimum of three 2.5-second epochs was selected, and these were used to calculate the absolute power of delta, theta, and alpha activity. For the present report, only the data from the midline electrodes (Cz, Pz, and Oz) are presented. Unlike the performance, oculomotor, and mood data, it was not possible to transform the EEG results into percent-change-from-baseline scores because no EEG data were collected prior to the deprivation period. Each of the activity bands were analyzed separately in independent two-way ANOVAs for time (2225, 0325, 0825, 1325, and 1825) and eyes (open and closed).

Delta activity. The analysis of absolute power within the 1.5-3.5 Hz range indicated a time-by-eyes interaction at Cz ($F(2.43,21.89)=3.31$, $p=.0473$) and at Pz ($F(2.20,19.78)=3.32$,

p=.0533). Analysis of simple effects indicated there were differences among the five testing times under the eyes-closed condition (but not the eyes-open condition) at both electrodes (p=.0182 for Cz, and p=.0040 for Pz). Trend analyses indicated that the differences were due to an overall linear increase in the amount of delta activity throughout the deprivation period ($p<.05$) as well as a cubic trend ($p<.05$). As shown in figure 7, slow-wave activity under the eyes-closed condition at Cz and Pz not only increased as a function of sleep loss, but there also was a particularly sharp accentuation from 0320 to 1320 after which there was a slight reduction from 1320 to 1820. In addition, the most noticeable difference between the two eye-closure conditions was at 1320 ($p<.05$). These findings generally agree with the significant time main effects found at Cz ($F(3.13,28.19)=3.56$, p=.0251), Pz ($F(3.70,33.34)=5.88$, p=.0014), and Oz ($F(2.42,21.77)=5.78$, p=.0069), all of which were associated with a linear increase in delta activity ($p<.05$) often punctuated by a cubic trend ($p\leq.06$) similar to the one described above. In addition, there were significant eyes main effects at all three recording sites which were due to more delta under eyes closed than under eyes open: Cz ($F(1,9)=9.15$, p=.0144); Pz ($F(1,9)=9.00$, p=.0150); and Oz ($F(1,9)=7.07$, p=.0261).

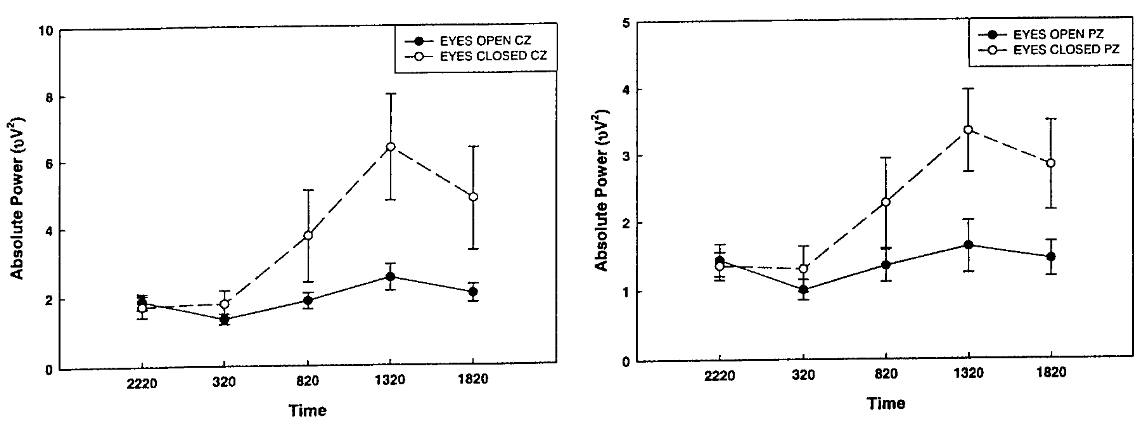


Figure 7. The effects of sleep deprivation and eye closure on delta activity at Cz (left) and Pz (right).

Theta activity. The analysis of absolute power within the 3.5-8.0 Hz range revealed a time-by-eyes interaction at Cz ($F(2.93,26.38)=3.04$, $p=.0478$), but not at the other two recording sites. Analysis of simple effects indicated differences across the testing times at eyes closed and at eyes open, but as shown in figure 8, the changes were most pronounced under eyes closed (where there was significantly more theta at 1320). Trend analyses identified linear and cubic deprivation effects ($p<.05$) under both eye-closure conditions. These were the result of 1) an overall deprivation-related increase in theta activity from the first to the last test session, and 2) an especially noticeable accentuation of theta at 1320, after which there was a decrease. This general pattern of results was consistent with the identification of a significant time main effect for Cz ($F(2.45,22.02)=3.89$, $p=.0289$) and Oz ($F(3.32,29.89)=5.46$, $p=.0032$), both of which were associated with significant linear and cubic trends ($p<.05$) similar to the ones described above. In addition, there were eyes main effects at Cz ($F(1,9)=5.12$, $p=.0499$) and Oz ($F(1,9)=6.38$, $p=.0325$) because of the presence of more theta under eyes closed than under eyes open.

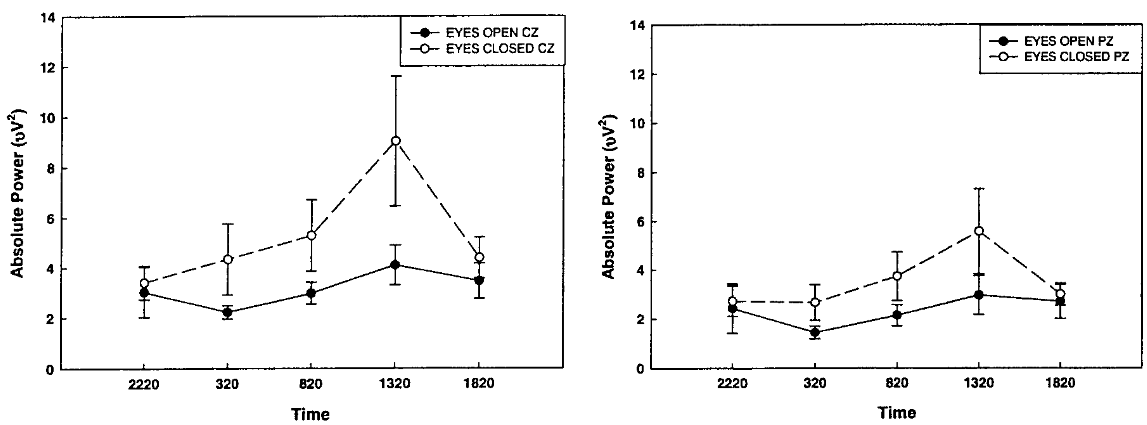


Figure 8. The effects of sleep deprivation and eye closure on theta activity at Cz (left) and Pz (right).

Alpha activity. The analysis of alpha power (8.0-13.0 Hz activity) indicated there were no statistically-significant effects associated with the testing times. In addition, eye-closure

exerted surprisingly little impact on the amount of alpha activity other than the marginally-significant ($p=.0603$) increase seen under the eyes-closed condition at Cz. Although there was a similarly-marginal time-by-eyes interaction at Cz ($p=.0670$), this effect also was not particularly robust. The means of the eyes-open condition at each electrode ranged from approximately 3.3 to 3.9 microvolts squared, while the means of the eyes-closed condition ranged from 5.3 to 6.4 microvolts squared; however, the variability was apparently too large for this difference to be considered "beyond chance." The polysomnographer who evaluated the EEG data for all of the volunteers indicated there were several participants who simply did not produce noticeable amounts of alpha activity under any circumstance.

Profile of Mood States

The data collected on the POMS yielded six factor scores for each iteration of this test. The factor scores reflected self-ratings of anger/hostility, tension/anxiety, depression/dejection, vigor/activity, fatigue/inertia, and confusion/bewilderment. Prior to analysis, these data were transformed into percent-change-from-baseline scores (the 2110 test on the training day was used as the baseline for the POMS). Afterwards, each set of factor scores was analyzed separately in six independent one-way ANOVAs to determine whether there were differences across the various deprivation testing times (2230, 0330, 0830, 1330, and 1830).

Tension/anxiety. The one-way ANOVA of tension/anxiety scores indicated there were no significant differences in the percentage of change from baseline on this scale.

Depression/dejection. The one-way ANOVA of depression/dejection scores revealed a significant overall time effect ($F(2.59,23.35)=3.97$, $p=.0243$), and subsequent trend analyses showed that this was due to the presence of a significant linear trend ($p<.05$) as well as a marginally-significant quadratic trend ($p=.0593$). As can be seen in figure 9, depression

scores increased in an overall way from the first to the last test session, and this was augmented by a particularly pronounced increase at the 0830 and 1330 testing times.

Anger/hostility. The analysis of percent-change scores on the anger/hostility scale indicated there were no differences on this scale as a function of sleep deprivation (testing time).

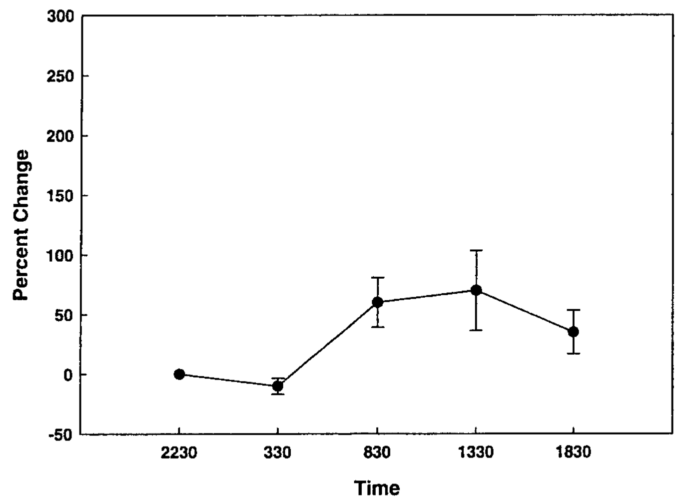


Figure 9. The effects of continuous wakefulness on POMS depression/dejection scores (expressed as percentage of change from baseline).

Vigor/activity. The one-way ANOVA of vigor/activity scores revealed a significant time main effect as a result of continuous wakefulness ($F(3.06,27.54)=13.74$, $p<.0001$). Trend analyses indicated this was due to an overall linear decrement in vigor levels ($p<.05$) as well as a quadratic change ($p<.05$) in self-ratings on this scale. As shown in figure 10, there was an especially large reduction in the vigor/activity scale (relative to baseline) at the 0830 test time, and this mood decrement remained at 1330 and 1830.

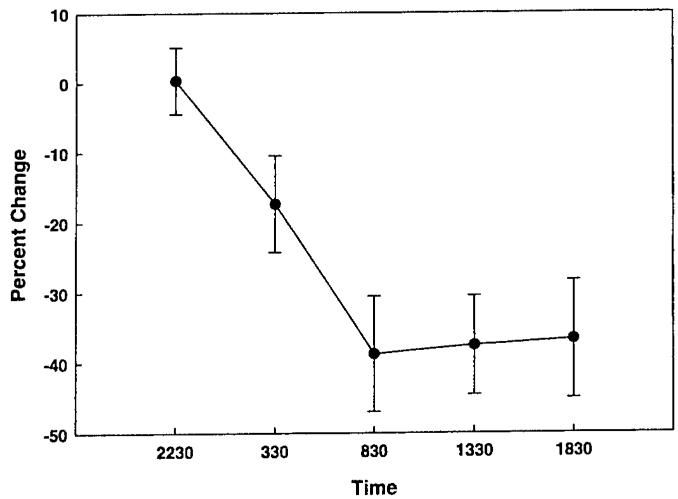


Figure 10. The effects of continuous wakefulness on POMS vigor/activity scores (expressed as percentage of change from baseline).

Fatigue/inertia. The analysis of the fatigue/inertia scale also revealed significant sleep-deprivation effects across the various testing times ($F(1.44,12.92)=11.77$, $p=.0024$). As was the case with the depression and vigor scales, subsequent analyses indicated the existence of both linear ($p<.05$) and quadratic trends ($p<.05$) in the data. Figure 11 shows there was an overall increase in percent-change-from-baseline fatigue scores from the beginning to the end of the deprivation period as well as a particularly large accentuation at 0830 and 1330.

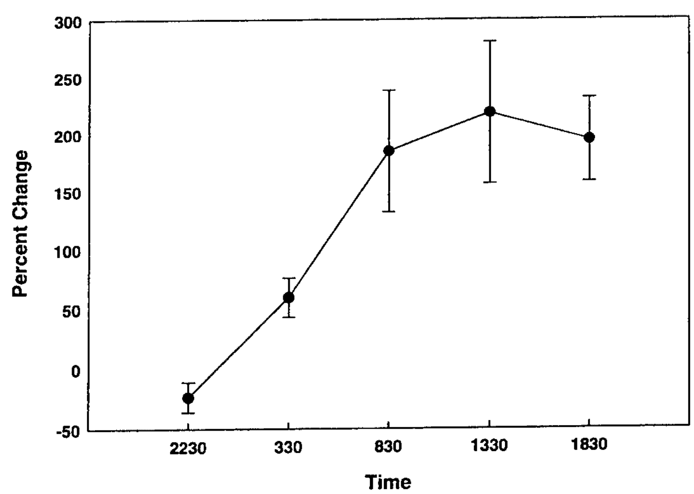


Figure 11. The effects of continuous wakefulness on POMS fatigue/inertia scores (expressed as percentage of change from baseline).

Confusion/bewilderment. The one-way ANOVA of scores from the confusion/bewilderment scale revealed a statistically-significant difference among the scores during the test sessions across the deprivation period ($F(3.16,28.43)=19.31$, $p=.0002$). Again, there were significant linear ($p<.05$) and quadratic trends ($p<.05$). As can be seen in figure 12, there was an overall increase in confusion ratings as well as an especially marked elevation (relative to baseline) at 0830. In addition, self-rated confusion remained high throughout the remainder of the deprivation period (i.e., at 1330 and 1830).

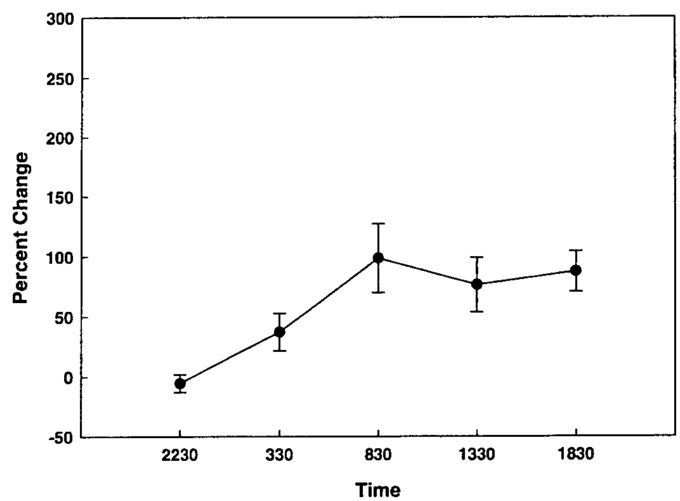


Figure 12. The effects of continuous wakefulness on POMS confusion/bewilderment scores (expressed as percentage of change from baseline).

Visual Analog Scales

There were a total of eight different VAS measures collected during each iteration of this test (VAS testing immediately followed the POMS administration). The VAS scores reflected self-ratings of alertness, anxiety, energy, confusion, irritability, jitteriness, sleepiness, and talkativeness. Prior to analysis, these data were transformed into percent-change-from-baseline scores, using the 2110 test on the previous training day as the baseline. Afterwards, each VAS measure was analyzed in separate one-way ANOVAs to determine

whether there were differences across the various sleep-deprivation testing times (2230, 0330, 0830, 1330, and 1830) on each scale.

Alertness. There was a significant difference in self-rated alertness as a function of sleep deprivation ($F(3.61,32.48)=13.30$, $p<.0001$) which was due to an overall linear decline in the scores from 2230 to 1830 (the linear trend was significant at $p<.05$) as well as an especially sharp decrease in alertness from 2230 to 0830 (the quadratic trend was significant at $p<.05$). Figure 13 shows that, following the substantial reduction in VAS alertness at 0830, there also was a slight recovery at 1330 and 1830; however, the ratings remained far below the ones observed at the outset of the deprivation period.

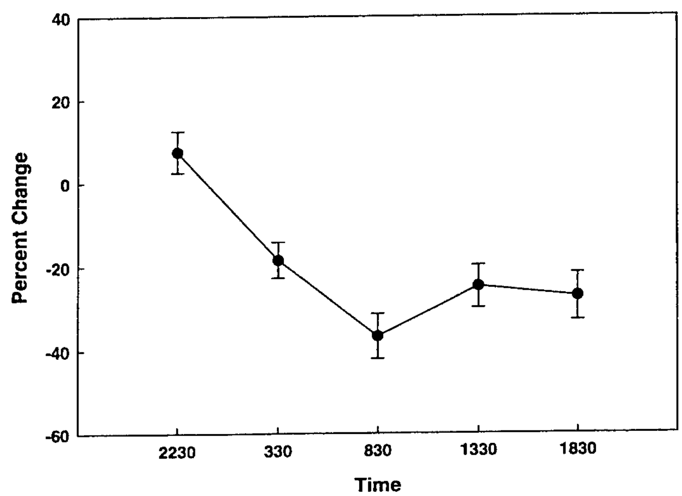


Figure 13. The effects of continuous wakefulness on VAS alertness scores (expressed as percentage of change from baseline).

Energy. The energy ratings revealed differences very similar to those found with alertness. First, there was a time main effect as a function of sleep deprivation ($F(3.40,30.62)=14.91$, $p<.0001$). Second, there was an overall reduction in energy ratings from 2230 to 1830 (significant linear trend at $p<.05$). Third, there was a quadratic effect ($p<.05$) due to the sharp drop in ratings at 0830 followed by a slight recovery at 1330 before energy ratings once again deteriorated at 1830 (see figure 14).

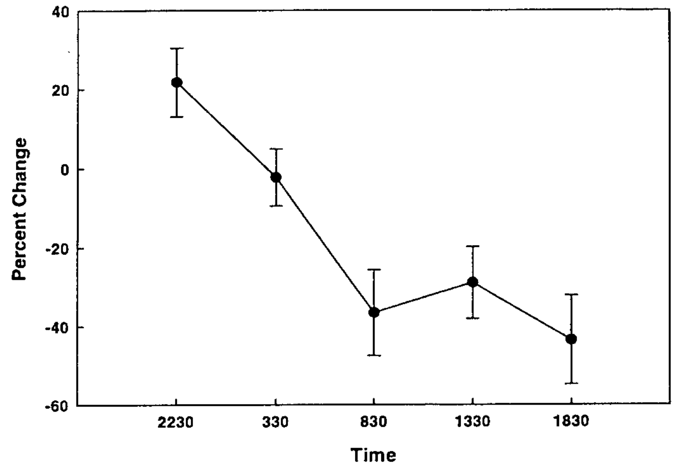


Figure 14. The effects of continuous wakefulness on VAS energy scores (expressed as percentage of change from baseline).

Anxiety. The VAS anxiety ratings were found to be unaffected by sleep deprivation in this study.

Irritability. VAS ratings of irritability were likewise unaffected by sleep loss. As was the case with VAS anxiety, there was no significant main effect on the time factor.

Jitteriness. Self-ratings on the jitteriness scale also did not change across the various testing sessions. Thus, the F-117 pilots tested here did not feel a change in their level of "nervous stimulation" as a function of sleep deprivation.

Sleepiness. VAS ratings on the sleepiness dimension surprisingly were not sensitive to the sleep loss to which the pilots were exposed despite the fact that alertness ratings were affected by the lengthy period of continuous wakefulness. The probability level on the time main effect for the sleepiness percent-change scores was only $p=0.093$.

Confidence. Self-ratings on confidence revealed a significant time effect ($F(2.95,26.53)=15.34$, $p<.0001$). Trend analysis indicated this was due to a decline in VAS confidence ratings from 2230 to 1830 ($p<.05$). Figure 15 graphically depicts this change.

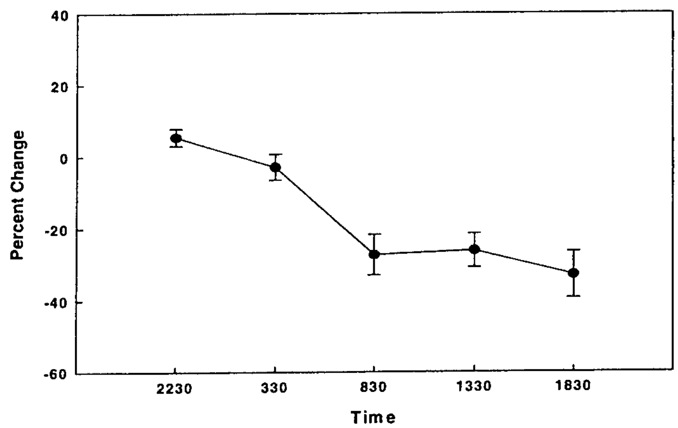


Figure 15. The effects of continuous wakefulness on VAS confidence scores (expressed as percentage of change from baseline).

Talkativeness. Self-ratings on VAS talkativeness were unaffected by the level of the sleep deprivation to which the participants were exposed in this investigation. The significance level for the overall analysis of these data was only p=.075.

In-flight (simulator) EEG

In addition to the resting EEG that was conducted under standard laboratory conditions, EEG data were collected during the simulator flights while the participants were flying the maneuvers discussed earlier. As was the case with the resting EEGs, data from the in-flight EEGs were not transformed into percentage-of-change-from-baseline values because there were no EEG data collected on the pre-deprivation training day. For the present report, data from the in-flight resting eyes-open EEG (immediately prior to the start of the flight profile), the data collected during the third straight-and-level, and the data collected during the fifth straight-and-level were examined. For the sake of clarity and brevity, these three EEGs will be referred to as Resting EEG, SL-3 EEG, and SL-5 EEG. Data collected during the remainder of the maneuvers will be the subject of an upcoming report.

Data analysis consisted of a series of 2-way ANOVAs for deprivation-test time (2300, 0400, 0900, 1400, and 1900) and EEG section (Rest, SL-3 EEG, and SL-5 EEG). Data from

Cz, Pz, and Oz were included. A separate ANOVA was performed for each of the EEG bands of interest (delta, theta, and alpha) as was the case for the EEGs collected under standard laboratory conditions.

In-flight delta activity. Analysis of slow-wave EEG activity in the 1.5-3.5 Hz range revealed a time-by-section interaction at Oz ($F(2.72,24.48)=3.01$, $p=.0537$) which simple effects indicated was attributable to differences among the sections at 0900 ($p<.05$) with a similar marginal effect at 0400 ($p<.07$). As shown in figure 16, there were larger reductions in delta activity from the resting EEG to the SL-3 EEG at 0900 than at the other times. Also, while not significant, the data during SL-3 seem to show an elevation in slow-wave EEG at 0400 and 0900.

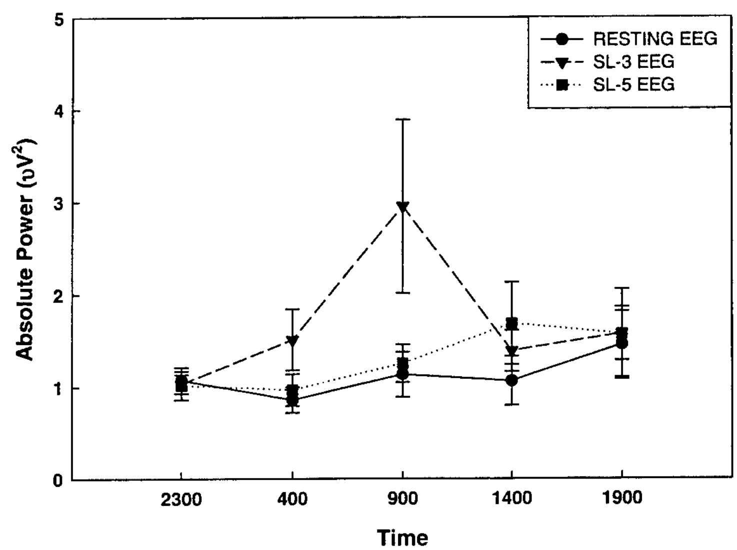


Figure 16. The effects of simulator-flight section (resting vs. SL-3 vs. SL-5) and time on delta activity at Oz.

In addition to the interaction, there were main effects on the section factor at Pz ($F(2,18)=6.92$, $p=.0059$) and Oz ($F(2,18)=5.40$, $p=.0146$); and a main effect on the time factor at Cz ($F(4,36)=5.88$, $p=.0010$). The section main effects were generally due to a substantial increase in the amount of delta activity ($p<.05$) during SL-3 EEG as opposed to the resting EEG, and to more delta at SL-3 EEG compared to SL-5 EEG (the comparison

between SL-5 and SL-3 was only marginally-significant at Oz). There effects are depicted in figure 17. The time main effect, also shown in figure 17, was a result of a deprivation-related linear increase in delta as a function of sleep deprivation (p<.05).

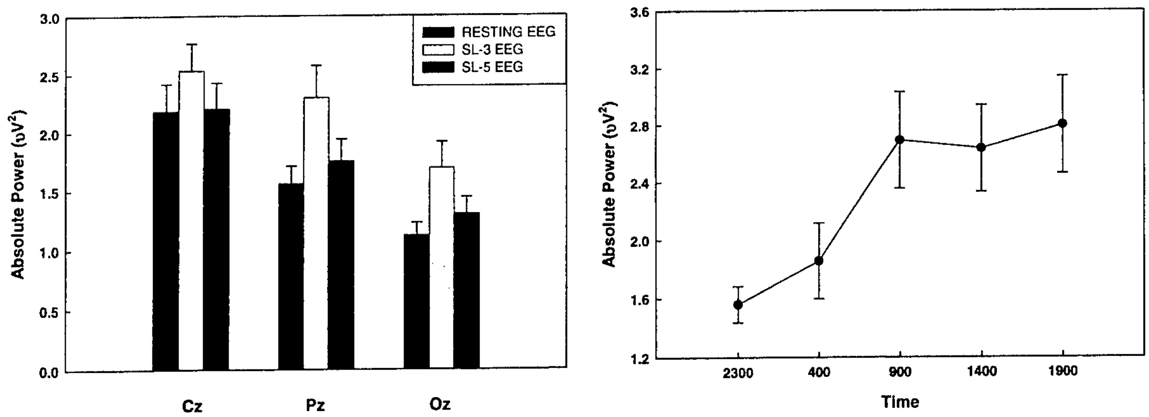


Figure 17. The effects of simulator-flight section (resting vs. SL-3 vs. SL-5) on delta activity at Pz and Oz (left), and the effects of time on delta activity at Cz (right).

<u>In-flight theta activity</u>. The analysis of theta activity (3.5-8.0 Hz) revealed only a single marginally-significant time main effect at Cz (p=.0755). The tendencies observed across the deprivation period are depicted in figure 18. There were no significant interactions or other effects.

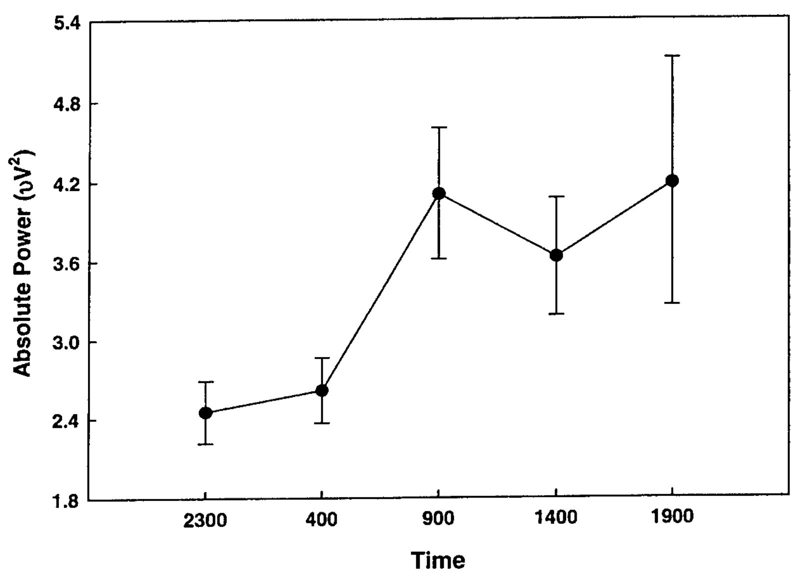
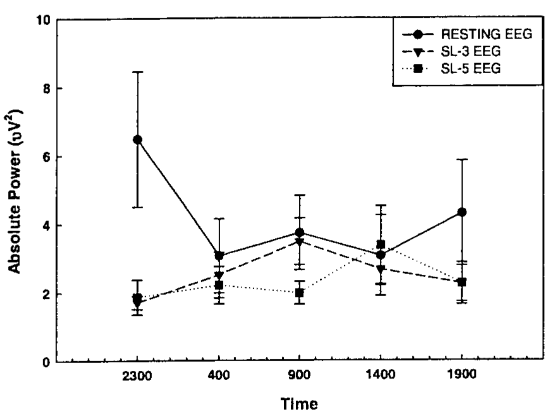


Figure 18. The non-significant time effect on theta activity at Cz.

In-flight alpha activity. The ANOVA on absolute EEG alpha power (8.0-13.0 Hz) indicated a marginally-significant time-by-section interaction, but in this case at Pz ($p=.0698$). The non-significant effect is shown below. In addition, there was a significant section main effect at this same recording site ($F(1.34,12.07)=5.34$, $p=.0314$). The section main effect was due to a substantial reduction in the amount of alpha activity from the resting period to the time during which participants flew the third and fifth straight-and-levels ($p<.05$), as shown in figure 19.



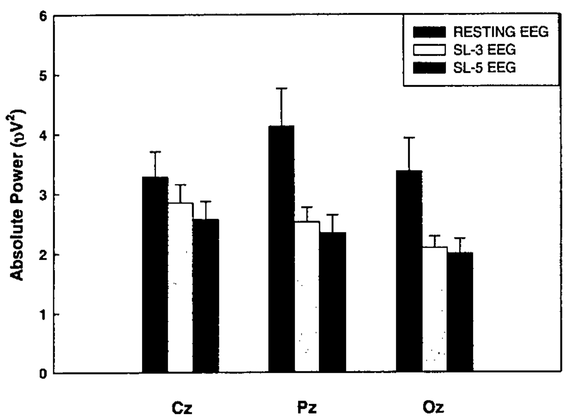


Figure 19. The non-significant time-by-section interaction on EEG alpha during the flights (left), and the significant differences among the recording sections (right).

Flight Performance

Prior to analysis, the data collected from each of the simulator flights were transformed into percentage-of-change-from-baseline accuracy scores by scoring each deprivation flight relative to the last flight on the previous training day (as described earlier). These percent-change data were first analyzed separately for each maneuver and then combined into a single overall analysis that included all of the maneuvers together. The separate analyses for the majority of the flight maneuvers consisted of one-way ANOVAs across the different testing times (2300, 0400, 0900, 1400, and 1900). However, the two right 360° turns and the five straight-and-levels were analyzed with two-way ANOVAs because there was more than

one iteration of each of these in the flight profile. Thus, for the right turns and the straight and levels, there were two design factors—maneuver iteration and testing time.

As detailed below, there were significant fatigue-related differences on all but one of the flight maneuvers when each was analyzed separately. The combined analysis confirmed an overall reduction in flight-performance skill as a function of sleep deprivation.

Straight climb. Analysis of the straight climb from 11,000 to 13,000 feet revealed that there was only a marginally-significant change in performance from the 2300-hour flight to the 1900-hour flight on the following day (p=.0666). As can be seen in figure 20, there was a tendency towards degraded performance at 0900 and 1400; however, the variability in the data prevented detection of an overall statistically-significant change in performance. Nevertheless, the data are presented for inspection since the combined analysis in which all maneuvers were considered together revealed an overall consistent change across each one.

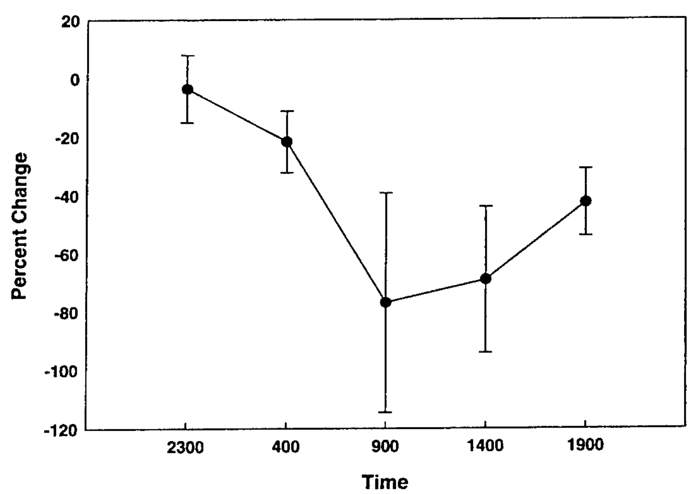


Figure 20. The marginally-significant effects of flight-testing times on the accuracy with which the straight climb was performed (expressed as percent change from baseline).

Straight descent. Percent-change scores from the maneuver in which participants descended from 15,000 to 13,000 feet on a consistent heading of 345° indicated a statistically-significant difference among the 5 flights conducted during the sleep-deprivation

period (F(3.67,33.03)=3.14, p=.0304). Follow-up tests revealed that there was an overall decline in performance from the first to the last flight (the linear trend was significant at p=.05), and that there was steep drop in performance accuracy from 2300 to 1400 which was followed by a slight recovery at 1900 (the quadratic trend was significant at p<.05). Figure 21 graphically depicts the changes that were observed.

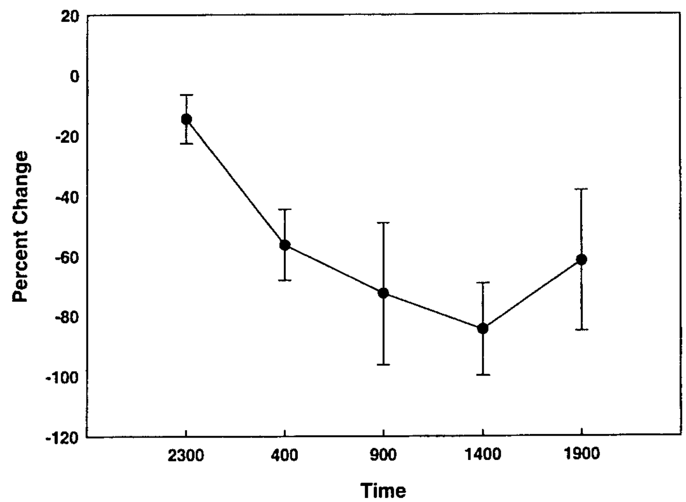


Figure 21. The effects of flight-testing times on the accuracy with which the straight descent was performed (expressed as percent change from baseline).

Left 720° turn. The data from the 720° left turn at 15,000 feet revealed a significant difference among the 5 flights (F(3.66,32.97)=7.38, p=.0003). Subsequent trend analysis indicated this was due to the presence of both a linear (p<.05) and a quadratic effect (p<.05). As is evident from the data depicted in figure 22, there was an overall reduction in performance accuracy from the first to the last deprivation flight (the linear effect) that was punctuated by particularly degraded performance at 0900 and 1400 (the quadratic effect).

Left climbing turn. Analysis of the data from the maneuver in which participants were required to climb from 10,000 to 15,000 feet while simultaneously executing a 540° turn indicated a significant change in performance accuracy across the 5 deprivation flights (F(2.91,26.17)=7.32, p=.0011). Once again, trend analyses revealed an overall degradation

in the performance of this maneuver from the first deprivation flight at 2300 hours to the last deprivation flight at 1900 hours the next day (the linear trend was significant at $p=.05$). The overall decline in performance was complicated by the fact that flight skills were especially impaired at the 0900 and 1400 flights (the quadratic trend was significant at $p<.05$). Figure 23 graphically depicts the changes that were observed.

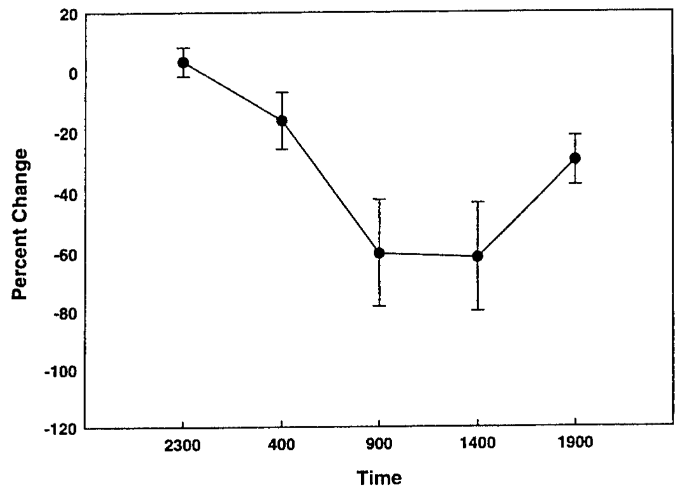


Figure 22. The effects of flight-testing times on the accuracy with which the 720° left turn was performed (expressed as percent change from baseline).

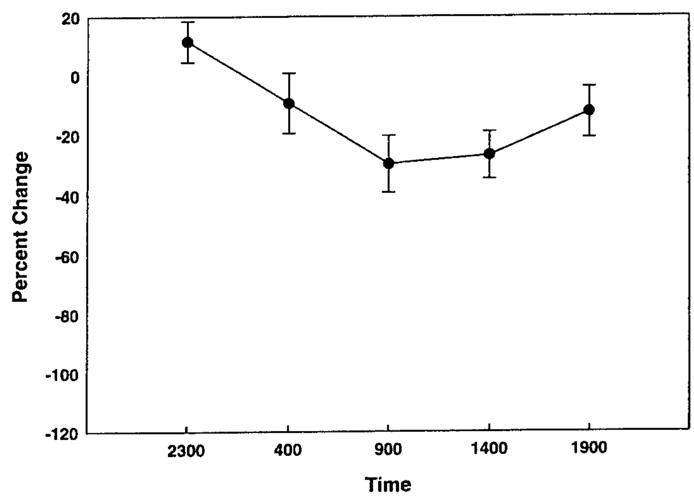


Figure 23. The effects of flight-testing times on the accuracy with which the left climbing turn was performed (expressed as percent change from baseline).

<u>Left 360° turn</u>. Analysis of the 360° left turn that was performed at an altitude of 11,000 feet revealed a significant change in accuracy across the 5 deprivation flights

(F(3.08,27.75)=7.21, p=.0009). Subsequent trend analysis revealed the presence of an overall linear reduction in performance ($p<.05$) as well as a quadratic effect attributable to especially poor performance at 0900 and 1400 ($p<.05$). As shown in figure 24, the decline in flight-control accuracy was fairly consistent from the 2300-hour flight through the 1400-hour flight, but there was a slight "end spurt" improvement at 1900 (although performance did not recover to the level that was observed at the outset of the deprivation period).

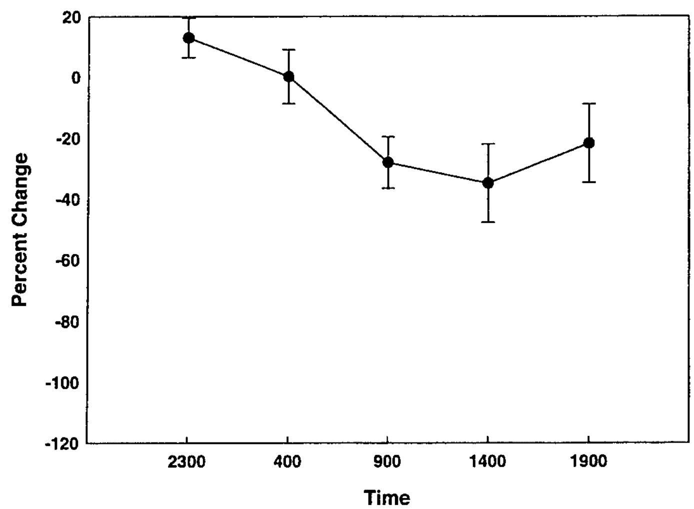


Figure 24. The effects of flight-testing times on the accuracy with which the 360° left turn was performed (expressed as percent change from baseline).

Right descending turn. Analysis of the maneuver in which participants were required to descend from 13,000 to 10,000 feet while simultaneously executing a 360° turn also indicated a significant change in performance accuracy across the various deprivation flights (F(4,36)=5.46, p=.0015). Trend analysis indicated this was due to an overall reduction in performance accuracy from the first to the last deprivation flight (the linear trend was significant at $p<.05$) as well as a particularly-noticeable degradation at 0900 (the quadratic trend was significant at $p<.05$). After 0900, performance remained degraded until the end of the deprivation period. Figure 25 graphically illustrates these effects.

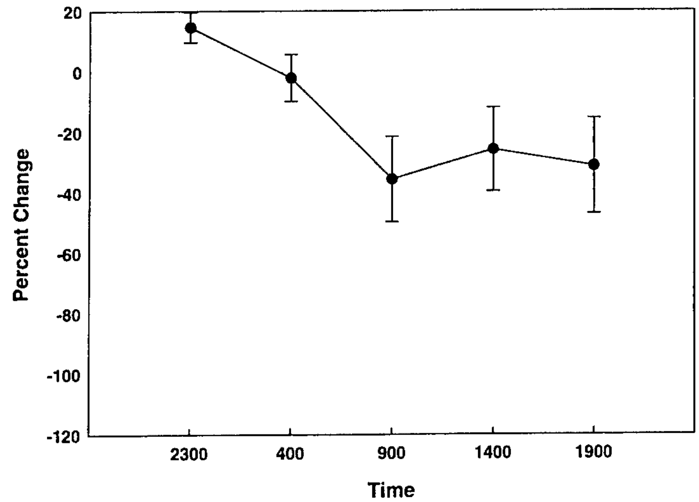


Figure 25. The effects of flight-testing times on the accuracy with which the right descending turn was performed (expressed as percent change from baseline).

Right 360° turn. The combined analysis of the first 360° right turn (that was performed at an altitude of 11,000 feet) and the second 360° right turn (that was flown at an altitude of 15,000 feet) revealed that there were no differences between the two turns and no differences in the turns' susceptibility to fatigue effects (neither the maneuver main effect nor the maneuver-by-time interaction was significant). However, there was a significant change in overall accuracy (with both of the turns collapsed) across the five deprivation flights ($F(4,36)=4.23$, $p=.0066$). Subsequent trend analyses revealed the presence of a linear change in performance ($p<.05$) as well as a marginal quadratic effect ($p=.06$). As shown in figure 26, these occurred because there was an overall decline in flight-control accuracy (relative to baseline) from the first deprivation flight at 2300 to the last flight at 1900 the next day, with an especially steep decline at 0900.

Straight and levels. The two-way ANOVA on all five straight-and-levels combined revealed that there were no overall differences among the various straight and levels (i.e., no maneuver main effect), and there was no maneuver-by-time interaction which would have indicated differential sensitivity to the effects of fatigue. There was, however, a significant

time main effect (F(2.83,25.50)=5.29, p=.0063) which was subsequently found to be a result of a linear decline in overall performance (p<.05) in conjunction with significant quadratic (p=.05) and cubic (p=.05) trends. As shown in figure 27, there was a steady decline in flight-performance accuracy from the 2300 through the 1400 flight with a slight "end spurt" during the final flight of the deprivation period.

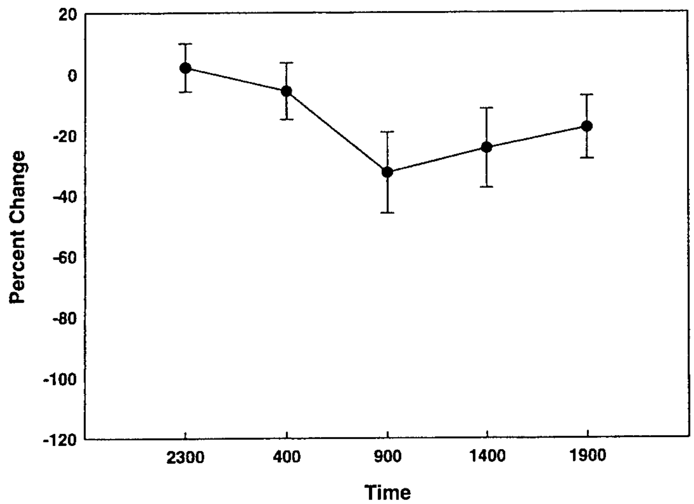


Figure 26. The effects of flight-testing times on the accuracy with which the right 360° turn was performed (expressed as percent change from baseline).

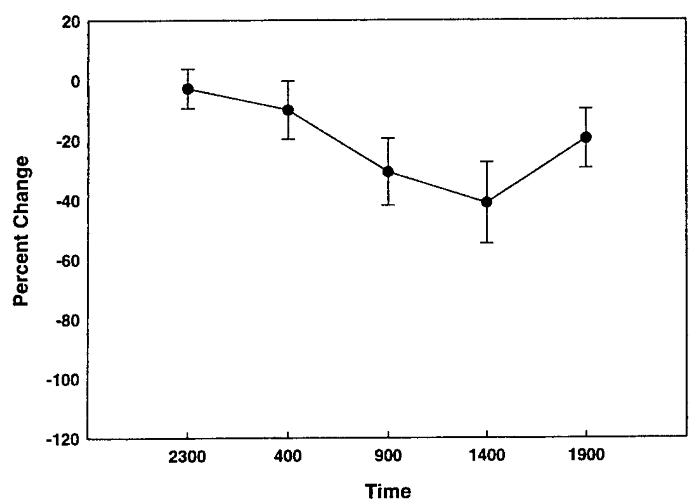


Figure 27. The effects of flight-testing times on the accuracy with which the right 360° turn was performed (expressed as percent change from baseline).

Composite flight performance. The two-way ANOVA that examined the impact of both testing time and flight maneuver on basic piloting skill across all of the flight maneuvers

considered together confirmed what was observed in the individual maneuver analyses. Specifically, there was a time main effect ($F(2.44,21.93)=10.72$, $p=.0003$) which subsequent analyses revealed to be a function of significant linear ($p<.05$), quadratic ($p<.05$), and cubic ($p=.05$) trends. As can be seen in figure 28, flight performance degraded from 2300 to 0900, remained consistently poor from 0900 to 1400, and then recovered slightly (although not to pre-deprivation levels) at 1900. In addition to these effects, the two-way ANOVA indicated there was an overall difference among the individual maneuvers ($F(4.79,43.10)=3.83$, $p=.0064$). This effect is not particularly surprising since some maneuvers may have been more difficult to perform than others; however, such a finding is not of interest since the time-by-maneuver interaction was not significant, suggesting that none of the maneuvers were more sensitive to the effects of fatigue than the others.

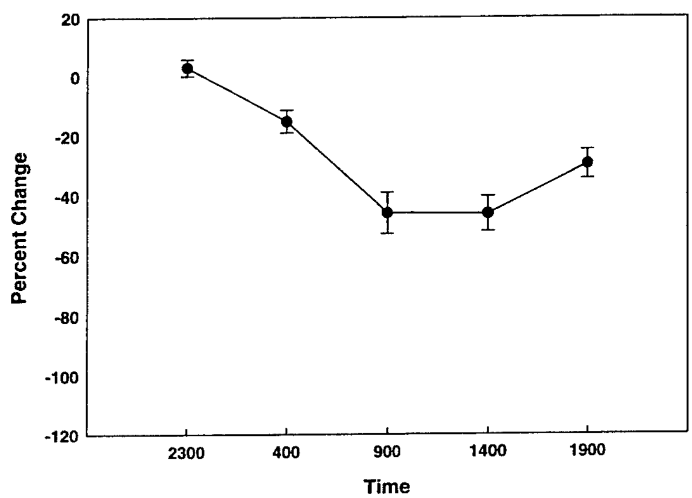


Figure 28. The effects of flight-testing times on the accuracy with which all flight maneuvers (analyzed together) were performed.

Predictive Relationships between Flight Performance and Select Measures

The correlations between the composite flight scores from each session and select non-flight measures from the same session were calculated on a individual-by-individual basis and then the 10 correlations obtained for each pair of measures were averaged (across all 10

participants). As stated earlier in the data-analysis section, the non-flight variables consisted of saccadic velocity, resting (laboratory) eyes-closed EEG theta activity at Cz, resting (laboratory) eyes-open EEG theta activity at Cz, MATB tracking errors, POMS fatigue, ANAM-mathematical-processing throughput, VAS alertness, and the in-flight EEG data recorded from Cz during the resting-eyes-open, SL-3, and SL-5 sections of the flights. Except for the EEG data, percent-change measures were used in each case.

The results of the participant-by-participant correlations are shown below in table 5, and the results of the averaged correlations are shown in table 6. For the present report, correlations of at least +0.59 or –0.59 were considered indicative of a noteworthy correspondence between two variables.

Note that the participant-by-participant correlations indicated substantial differences in the strength of the relationships among the various variables as a function of the measure under consideration and as a function of individual differences. This accounts for the fact that only 43 percent of the coefficients attained a magnitude large enough to be considered important. Nonetheless, when examining the relationships measure-by-measure, it became clear that some measures possessed a greater potential for predicting flight performance than others. Of the physiological measures, the amount of theta activity during the last straight-and-level was most closely related to flight performance as indicated by relatively high correlations in six of the ten participants. Furthermore, each of the correlations were in the "intuitively correct" direction (i.e., increased theta activity was associated with decreased performance accuracy). Of the two self-report measures, the scores on the POMS fatigue scale were most closely related to performance. Correlations of .59 or greater were found within eight of the ten participants, and once again, all of the relationships were in the expected direction (i.e., increased fatigue was correlated with reduced accuracy). Of the

cognitive measures (MATB tracking and ANAM mathematical processing), MATB tracking was the better indicator of subsequent flight performance, at least for most of the participants. However, despite the fact that strong relationships were observed within eight out of ten volunteers, in one of these cases, the relationship was unexpectedly positive (suggesting that increased MATB errors were predictive of improved flight performance) rather than negative (suggesting that increased tracking errors were indicative of impaired flight performance). The remaining variables-- saccadic-velocity scores, resting EEG data (during both the laboratory and in-flight conditions), the EEG data collected during the performance of straight-and-level 3, and self ratings of alertness on the VAS--were not consistently associated with changes in flight skills across the sample (although in a few participants, these measures were very good at suggesting upcoming performance impairments).

Table 5. Correlations between Composite Flight Score per Subject and Various Performance/Physiological Data

Subject	Seccadic Velocity	Resting Cz Theta	MATB Tracking Errors	POMS Fatigue	ANAM Math Throughput	VAS Alertness	Preflight Cz Theta Resting	Flight Cz Theta SL3	Flight Cz Theta SL5
01	0.43	-0.15	-0.98	-0.89	-0.28	0.74	0.67	-0.55	-0.75
02	0.76	0.51	-0.89	-0.74	0.77	0.59	-0.28	0.03	-0.60
03	0.07	0.59	0.64	-0.93	-0.27	0.92	0.06	-0.06	0.39
04	0.85	-0.95	-0.92	-0.84	0.29	0.09	-0.98	-0.29	-0.88
05	-0.50	-0.81	-0.79	-0.99	-0.05	0.57	0.53	-0.14	-0.29
06	0.79	-0.52	-0.45	0.07	0.17	0.48	0.00	-0.97	-0.77
07	0.14	-0.40	-0.71	-0.77	0.81	0.59	-0.27	-0.01	-0.65
08	0.36	-0.43	-0.95	-0.53	0.46	0.88	-0.69	0.00	-0.60
09	-0.11	-0.88	-0.96	-0.87	0.89	0.03	-0.49	-0.25	-0.23
10	0.83	-0.66	-0.26	-0.73	0.52	0.56	0.58	-0.52	-0.21

The average overall correlations (across all 10 participants) generally agreed with what was noted above. As shown in table 6, many of the promising measures failed to demonstrate a consistent relationship to overall flight performance, but of the cognitive measures, MATB tracking was most highly related to overall flight-performance accuracy,

and the direction of the relationship was intuitively correct (increased MATB errors were associated with degraded flight skills). ANAM throughput tended to be correlated with flight performance, but to a lesser extent than the MATB data. Among the physiological measures, the oculomotor measure and some of the EEG data showed a moderate degree of overall predictive utility in that flight-performance reductions tended to be associated with longer saccadic velocities as well as increased EEG theta activity (at least in the resting EEG and the one collected during the fifth straight-and-level), but the relationships were not as strong as had been expected.

Table 6. Average correlations between flight accuracy and potential predictive measures collapsed across all 10 participants.

	Saccadic Velocity	Resting CZ Theta	MATB Tracking Errors	POMS Fatigue	ANAM Math Throughput	VAS Alertness	Preflight CZ Theta Resting	Flight CZ Theta SL3	Flight CZ Theta SL5
Flight Score	0.36	-0.37	-0.63	-0.72	0.33	0.55	-0.09	-0.28	-0.46

In summary, there was evidence that some of the physiological indices, the non-flight performance measures, and a subset of the self-reported mood ratings could offer an indication about the accuracy of future flight performance, but with the exception of MATB tracking and POMS fatigue, the associations were relatively weak. However, it should be noted that these findings are based on bivariate correlations that were (by necessity) calculated on only a small set of observations (5 pairs each), and that many statisticians recommend the use of at least 100 pairs of scores for the determination of reliable correlation coefficients. Thus additional work remains to be done before definitive conclusions can be drawn. In the future, it may be possible to collect additional observations under similar circumstances, combine the data sets, and alleviate this concern.

DISCUSSION

This study objectively evaluated the impact of long hours of sustained wakefulness on the self-reported alertness, cognition, physiological arousal, and flight performance of rated military aviators. It is the only assessment of its type that has been conducted to specifically assess the impact of fatigue on current active-duty F-117 pilots. In addition, the investigation was designed to assess whether measures of physiological arousal (oculomotor data and EEG data) hold promise for predicting upcoming performance decrements. The results indicated a variety of operationally-relevant effects that individual aviators and their commanders may wish to consider prior to training for, scheduling, and actually conducting sustained operations.

The Overall Impact of Fatigue

First of all, there were noteworthy cognitive decrements attributable to sleep loss, although not every cognitive measure was adversely affected. Two of the eight performance measures from the MATB and two of the four measures from the ANAM mathematical processing task were severely degraded as a function of 37 hours of continuous wakefulness. On the MATB, both response times on the monitoring task (in which participants had to vigilantly monitor warning lights and dials) and the accuracy of tracking a target were progressively impaired across the deprivation period, with the onset of particularly noticeable degradations at 0730 (after being awake for 25.5 hours). The decrements persisted until the last test session, 10 hours later. However, the accuracy and speed with which the pilots responded to simulated radio calls, and their ability to monitor simulated fuel levels were unaffected. On the mathematical processing task, overall response time was degraded in a manner similar to what was observed on the MATB, and overall throughput (number of

correct responses per minute) was likewise impaired; however, response accuracy was not affected.

Generally speaking, these types of findings are consistent with the results from previous sleep-deprivation studies (such as Caldwell, Hall, and Erickson, 2002; and Caldwell and Ramspott, 1998). In addition, Wilkinson (1964) long ago reported that serial reaction time and vigilance tests were significantly degraded by sleep loss. When possible, sleepy people typically trade speed in order to preserve accuracy. However, when the task is machine paced (such as MATB tracking), this strategy becomes ineffective because slower performance inevitably leads to increased errors by decreasing the overall number of rapid, small control corrections. Since piloting is an externally-paced task, the types of reaction-time and tracking degradations found in this investigation suggest likely reductions in militarily-relevant performance in the operational aviation environment. Larger flight-path excursions should be expected as well as a tendency to under-correct targeting deviations when in manual tracking modes.

Second, there was clear evidence of decrements in central nervous system arousal as a function of increased operator sleepiness. Resting EEG activity (collected prior to the simulator flights) revealed a systematic slowing of brain activity across the deprivation period, with especially noticeable effects at the 1330 test session (after 31.5 hours awake), while oculomotor changes in half of the measured parameters (pupil-constriction speed and saccadic velocity) showed the greatest changes at approximately 0300 in the morning (after 21 hours awake). The "in-flight" EEG data (collected before "departure" while participants were seated in the simulator and then while flying two of the maneuvers) was not as sensitive to the presence of fatigue effects as the standard laboratory EEGs. In all probability, this was due to the fact that the participants were either preparing for the flight (prior to departure) or

making every effort to actively perform the required flight maneuvers (during the flight), and this tended to offset many of the short-term sleepiness effects that were more pronounced in the somewhat more soporific laboratory environment. Nonetheless, there was a pronounced increase in delta activity and a tendency for increased theta activity at Cz as a function of sleep loss. As was the case with the standard laboratory EEGs, the most pronounced slow-wave elevations occurred during the 1400 flight (after 32 hours awake), but there were also marked increases five hours earlier at 0900.

Past research has established that sleepiness and fatigue elevate the amount of slow-wave brain activity (Pigeau, et al., 1987), and increased theta (3.5-8.0 Hz activity) has been associated with generalized decrements on cognitive tasks (Belyavin and Wright, 1987) as well as reduced speed of responding to incoming stimuli (Ogilvie and Simons, 1992). Increased delta activity, especially in the parietal region (which is presumed to reflect deactivation of brain areas associated with attention and visual peripheral awareness (Mesulam, 1985)) logically would be expected to cause a variety of performance lapses. Both EEG-based investigations and brain-imaging studies (i.e., Thomas et al., 1998) have shown that this area of the brain is one area that is particularly deactivated during sleep deprivation, and in this study, parietal deactivation was suggested by increased slow-wave EEG activity recorded from Pz (the parietal midline electrode) as well as a similar effect at Cz (the vertex electrode).

Oculomotor parameters likewise have been shown to be sensitive to fatigue. Both constriction latency and saccadic velocity become slower as the level of sleep loss increases, with saccadic velocity being especially sensitive to sleep deprivation (Russo et al., 2003). Reduced saccadic velocity may reflect decreased motivation, reduced attention, and neuronal changes that undermine mental processing speed (Balkin et al., 2000).

Of course, these types of changes theoretically could lead to a variety of problems in operational contexts. Reduced brain activation (signaled by EEG, oculomotor, and other physiological changes) likely will degrade such functions as response speed, attention, memory, higher-level judgment, and vigilance in every aspect of the military aviation environment. These degradations should be expected to impair overall mission readiness.

Third, self-reported mood and alertness ratings revealed a number of adverse changes associated with sleep loss. POMS vigor declined substantially from the beginning to the end of the deprivation period, with the most noticeable reductions from 2200 to 0800 the next morning (although decrements persisted until the end of testing). Similar effects were seen in VAS measures of alertness, energy, and confidence. Meanwhile POMS ratings of depression, confusion, and fatigue all revealed an overall increase as a function of extended wakefulness, with the most striking effects occurring (once again) by the 26th hour of continuous wakefulness (although depression and fatigue tended to increase further through the 31st or 32nd hour). Needless to say, such fatigue-induced mood impairments can affect task motivation, increase the amount of perceived effort required to accomplish mission objectives, and interfere with interpersonal relationships and willingness to communicate (resulting in degraded crew coordination).

Fourth and most important, the objective flight-performance data indicated substantial decrements in the pilots' basic abilities to precisely maintain target headings, altitudes, airspeeds, bank angles, and vertical-velocities as a function of increased sleep loss. The consistently-observed pattern of effects reflected an overall decline in performance accuracy from the beginning to the end of the deprivation period with the most noticeable degradations occurring after 27 hours awake (although on some maneuvers, the flight that occurred after 32 hours of wakefulness was just as degraded as the one flown five hours earlier). At these

times in particular, the pilots tested in this study tended to become frustrated at their own inability to precisely maintain specific flight-path parameters, and several of them later indicated that they tended to relax their own precise standards of performance just for the sake of getting through the flight profile.

The timing and magnitude of the reductions in general flight skills are partially consistent with those found in previous studies (Caldwell et al., 1995; Caldwell et al., 1996; Caldwell and Caldwell, 1997), although the period of time in which performance was most impaired tended to be later by approximately three to five hours than what has been observed earlier. This initially curious change in the timing of the performance troughs (relative to previous studies with Army helicopter pilots) is likely due to the fact that the pilots tested in the present evaluation tended to be on a later work schedule than the pilots used in the Army studies. Many members of the present pool of participants were working a daily schedule that required an 0900 to 1000 report time, whereas the majority of the earlier Army participants were working a daily schedule which required an 0600-0700 report time. This scheduling difference was apparently responsible for the fact that the Army studies showed the poorest flight performance between 0500 and 1000 (after sleep deprivation) while the present study revealed the poorest flight performance between 0900 and 1500.

In any event, the present results confirm that performance in sustained operations is substantially affected by two primary factors: the amount of time since the last sleep period and the time of day according to the body's internal clock (Akerstedt and Folkard, 1997). Especially at times when the sleep pressure from numerous hours of continuous wakefulness combines with an increase in body-clock-driven (circadian) sleep pressure, alertness, vigilance, and performance will suffer.

Interestingly, a comparison of the flight effects with the results of the other assessments conducted in this study showed that: 1) the greatest performance and mood decrements occurred at similar points in time, and 2) the timing of the greatest decrements did not coincide with the presumed timing of the circadian trough (i.e., at 0300 or 0400 in the morning). Instead, the overall deterioration in functioning was later than most people (including many of the pilots tested in this investigation) often assume that it would be, and the decrement persisted for several hours longer than many would have predicted. This same pattern of effects was observed earlier in our studies with helicopter pilots (although here, the most impaired times were even later). It is noteworthy that several of our participants walked out of the simulator building around the time of sunrise explicitly for the purpose of deriving an alertness-enhancing benefit from sunlight exposure. Afterwards, they often verbally reported that because of their bright-light exposure, they felt more awake. However, neither the flight data nor the POMS/VAS-based self-report questionnaires supported this contention. Ultimately, flight performance remained substantially-degraded as did most other aspects of performance examined here, well into the daylight hours. Although no one crashed, the clear-cut loss of basic flight-control skill suggests that higher-level judgments, decisions, and other aspects of mission readiness would have suffered in the operational settings.

The potential operational significance of the observed findings may be greater than they at first appear when considered in light of the fact that our research procedures (which have identified several problems) are in a variety of ways more alerting than many of the work situations that exist in real-world fighter and bomber missions (except for actual combat and a few others). A primary example is the fact that the flight profile used in this investigation was relatively short in comparison to the flights that occur in the operational environment. In

fact, it took only one hour to complete each of the simulator-based assessments used here, and after each assessment, the participants had opportunities to stand up, walk around, converse with others, snack, and watch TV or play video games. Furthermore, during each of the flights the pilots were verbally cued about when to begin every one of the 15 maneuvers that were flown (a fact that in-and-itself provided several "wake-up" calls throughout each flight assessment). All of these factors no doubt enhanced subsequent alertness and performance, at least for a while. LeDuc et al. (1999) found that exercise temporarily decreased sleepiness in helicopter pilots, Neri et al. (2002) determined that even short hourly activity breaks were extremely beneficial for maintaining the alertness levels of long-haul commercial pilots, and Caldwell, Prazinko, and Caldwell (2003) documented the arousal-enhancing properties of changes in body posture. However, needless to say, the real-world, long-range, high-altitude sorties in which F-117 pilots are strapped into ejection seats, alone in the cockpit, for up to 18 straight hours are devoid of opportunities for such beneficial manipulations. This makes it a virtual certainty that any of the decrements observed in this study will be further compounded in real-world operations.

In summary, the fatigue-related degradations in cognitive, mood, and objectively-measured flight performance that occurred here as the result of 37 hours of continuous wakefulness provide direct evidence that actual operational readiness will suffer unless sleep/wake schedules and duty periods are properly managed, and/or pharmacological or non-pharmacological fatigue countermeasures are effectively employed in actual sustained operations. Despite opinions to the contrary, the degradations found here are common across many individuals, and no amount of training, professionalism, or motivation is likely to prevent their occurrence in the real world.

The Possibility of Predicting Degraded Performance

Having documented the negative impact of fatigue on pilot performance and alertness, the next topic of interest centers around whether or not it might be possible to predict operationally-relevant decrements based upon either objective physiological measures, the results of cognitive tests, or self-reported estimates of fatigue or alertness. If it is possible to make such predictions, the outcome could be used to precisely guide the implementation of countermeasures and/or to warn individual personnel about the increased likelihood of an upcoming performance failure.

The fact that the time courses of the oculomotor effects and some of the EEG changes were different from one another and different from the time course seen for the basic cognitive decrements and performance impairments is an interesting finding that does not bode well for the establishment of reliable one-to-one predictive relationships. In fact, the participant-by-participant bivariate correlations between composite flight performance and select EEG-theta results and oculomotor-saccadic-velocity data were relatively small (with absolute values of the best correlations ranging from approximately 0.3 to 0.5). However, it appears that a small, potentially-useful relationship may very well exist since reduced saccadic velocity tended to be associated with reduced flight accuracy. This coincides with the general observation (seen in the graphic depictions of the present data) that oculomotor differences occurred before flight-skill deteriorations were seen, suggesting that at least a subset of these physiological measures can offer insight into the potential for upcoming reductions in mental or performance capabilities. Russo et al. (2003) earlier reported that slower saccadic velocity was correlated with increased driving accidents and increased sleepiness, so the present results tend to support these earlier findings.

The EEG changes (from the laboratory resting EEG) indicated the onset of fatigue effects at about the same time that performance sharply declined, but peak EEG differences occurred later than peak degradations in basic cognitive skills or flight skills. Thus, while the EEG data were not highly correlated with flight performance, they seem useful for predicting the onset of an overall reduction in performance capacity, and they may be useful for estimating increases in the overall level of the homeostatic drive for sleep (since brief episodes of slow-wave EEG activity tend to become more frequent with increased sleep pressure). However, at least in this study, standard pre-flight EEG evaluations did not appear especially useful for making accurate, individualized, moment-to-moment predictions about subsequent flight-performance changes.

The EEG data collected during the actual performance of the flight maneuvers also did not show as strong an individualized relationship to fatigue-induced performance decrements as had been hoped. Although increased theta activity (one typical indication of increased sleepiness) did correlate with decreased performance accuracy in the -0.3 to -0.5 range, this did not approach the -0.7 to -0.8 range that was initially anticipated. However, the in-flight EEG data nonetheless seemed to be somewhat useful for making real-time indications of impairments in upcoming flight performance since the small correlations were at least in the "intuitively correct" direction. It is possible that workload factors may be responsible for clouding the interpretation of these data since some investigators (i.e., Sterman et al., 1987) have reported a positive relationship between theta activity and increased workload. Perhaps the perception of increased effort that often results from trying to perform tasks in the face of increased fatigue actually resulted in some of the same physiological changes that have been found to occur as a result of increased task difficulty. Additional analyses will be required to further explore these possible relationships, and to determine whether oculomotor or EEG

data can reliably be used to establish "fitness for duty." One possibility is that better flight-skill predictions might result from correlating flight performance with the physiological data from *four or more hours prior* to the simulator sessions as opposed to correlating this performance with the physiological data that was recorded *immediately before or during* each of the flights.

With regard to the non-physiological data, the time course of changes in cognitive performance and self-reported mood were similar to those we have observed in earlier studies (Caldwell et al., 2000; Caldwell, Hall, and Erickson, 2002), and correlational analyses of these data suggested that the cognitive data were more predictive of subsequent flight performance than were the other measures. Decrements in preflight MATB tracking corresponded very well to flight-skill degradations (the relationship was -0.63), and ANAM-mathematical-processing throughput corresponded to some degree with flight-performance accuracy (0.33). Both of these tests were sensitive to the impact of sleep deprivation, and surprisingly, at this point at least one of them (the MATB) seems to hold the greatest promise for predicting aspects of operational readiness.

The basic pattern of self-reported mood decrements also appeared to be basically consistent with the overall changes in both cognitive performance and flight performance. In fact, POMS fatigue ratings correlated with subsequent flight performance better than any other measure, with a correlation of -0.72. When fatigue ratings increased, flight accuracy decreased. In addition, while self-assessments of alertness were not as closely associated with subsequent performance as were the fatigue ratings, they nonetheless seemed to hold noteworthy predictive utility (the correlation was 0.55). These statistical findings coincided with the graphic depictions of the results which showed that the greatest increases in mood disturbances occurred after 26 hours of continuous wakefulness—the very times at which

cognitive skills and flight performance appeared to be most degraded. In some ways, this is surprising since it is often said that humans are not reliably able to predict the points at which they are at the greatest risk for uncontrolled sleep attacks or serious lapses in vigilance (Dinges, 1989). However, they do seem capable of knowing when they are generally becoming very sleepy and unresponsive (as indicated by Horne and Reyner, 1995), and this fact is confirmed by our present results. However, this does not necessarily mean that fatigued people are able to *consistently* anticipate substantial moment-to-moment deviations from assigned flight parameters. In fact, several of the participants in the present study were unable to accurately identify the times at which the most serious flight-performance decrements occurred.

Thus, while the majority of measures collected in this study were sensitive to the deleterious effects of extended wakefulness and circadian factors, and while they provided an overall indication of when fatigue-related decrements were likely to occur, the only strong overall relationships between non-flight data and actual flight skills were observed with psychomotor-tracking performance and self-rated fatigue. The physiological measures (oculomotor and EEG) did not show a sufficiently large one-to-one correspondence with the flight evaluations to allow accurate individualized predictions about the extent of upcoming flight-performance degradations. Nonetheless, the overall pattern of effects did provide an indication of the times at which personnel engaged in sustained operations are particularly at risk, and within specific individuals, the predictive relationships were quite strong. Perhaps the greatest challenge in the operational context will be to determine which subset of measures will be most useful for predicting the operational performance of the greatest number of personnel.

CONCLUSIONS

Mood, cognition, and physiology were adversely affected by 37 hours of continuous wakefulness, and these impairments degraded the objectively-measured ability to maintain basic flying skills. The overall complex of the present findings suggests that when pilots are exposed to a high degree of sleep loss (especially during circadian "low points"), 1) crew coordination will likely decline (due to the impact of poor psychological mood on interpersonal interactions), 2) the rapid processing of information will become more difficult (due to impaired attention), and 3) a wide array of basic performance capabilities will be degraded (due to compromised vigilance, poor situational awareness, and sluggish reaction time). Even well-practiced, long-standing, flight skills, in pilots who possess hundreds of flight hours, will be significantly impaired.

Although none of the pilots in this study crashed the simulator, the fact that they lost a substantial amount of flight precision in a research environment that is far less demanding than the actual operational context indicates that the more complex types of flight tasks demanded by more stressful real-world conditions are likely to be dangerously compromised after 22 hours or more hours of continuous wakefulness unless appropriate fatigue countermeasures are employed at the correct times. The general timing of the most serious operationally-relevant decrements following a day and night of sleep loss appears to be from approximately 0700 to approximately 1600 (in personnel who know they will be released from duty at the end of the day). However, it is not yet possible to make accurate individualized predictions about the exact timing and degree of outright performance failures despite the fact that some of the non-flight measures collected in this investigation hold promise for warning personnel about the possibility of future flight-related problems.

Along these lines, two critically important points were highlighted by the present study: 1) pilots cannot reliably determine the points in time at which their flight performance (and any other aspect of their functioning) is most impaired; and 2) they are often incorrect in assuming that the most dangerous performance period following many hours of continuous wakefulness is between 0300 and 0600 in the morning. In fact, despite beliefs to the contrary, alertness and cognition were not improved by the dawning of a new day, and in fact, even well-practiced flight performance continued to decline until as late as 1400-1500 in the afternoon (after 32-33 hours awake). Finally, although the majority of participants in this study were familiar with sleep deprivation based upon their own operational experiences, and despite the fact that they were highly motivated professional pilots, these factors did not protect them from fatigue-related impairments that are based on immutable physiological factors.

All of these points serve to further highlight the importance of training individuals and leaders about basic fatigue facts as well as educating all personnel about scientifically valid ways to combat fatigue in the operational environment. In addition, this study suggests it is valuable to objectively assess the potential benefits of specific alertness-enhancing countermeasures within the same context and within the same population that ultimately will rely upon these countermeasures in the operational environment. Whenever possible, data that will affect the lives of pilots should be collected from real pilots performing at least a subset of actual piloting tasks.

ACKNOWLEDGMENTS

The authors are grateful to the Command and Staff of the 49th Fighter Wing at Holloman Air Force Base for their unequivocal and exceptional support of this research. In particular, we would like to thank the Commanding General, Brigadier General James P. Hunt; as well

as the Vice Commander, Colonel C. G. C. Treadway; the 49th Operations Group Commander, Colonel Andrew W. Papp; the Commander of the 7th Combat Training Squadron, Lieutenant Colonel Angelo B. Eiland; and the Simulator Liason Officer, Major Richard Beavers; as well as every member responsible for the operation of the Link F-117A Weapons System Trainer including Mr. Dean Wardell, Mr. Laurence Chaky, Mr. Rodrigo Duran, Mr. Jeffrey Glaze, Ms. Kellie Glaze, Mr. Mark Krieger, Mr. Kerry McClannahan, Mr. Brian Miller, Mr. Noel Terry, Jr., Mr. Robert Unger, and Mr. Dennis Whyman, as well as Mr. Adolf Wolfer of the Training Systems Squadron. The exceptionally cooperative attitude and expert performance of all of these individuals ensured that the simulator was fully operational at all times, the flight data were flawlessly stored and downloaded, and that the research staff was provided with every measure of on-site assistance that could possibly have been required. In addition, we owe a sincere debt of gratitude to the F-117 instructor pilots who helped refine the research flight profile, trained members of the research staff to expertly structure and conduct each simulator test session, and generally supported the research staff during each volunteer's training flights. The contributions of Mr. James Brown, Mr. Richard Guidry, Mr. James Rials, Mr. Norman Schoening, and Mr. Jeffrey Ault, significantly enhanced the feasibility and operational relevance of this work. Besides these individuals, we would like to thank the Wing Scheduler, Ms. Irene Rivas for coordinating the simulator availability times, and the Base Flight Surgeon, Lieutenant Colonel Gordon Peters and his staff (most notably Technical Sergeant Bryan Cole) for providing indispensable medical support. We are indebted to those who provided the generous financial support for this project, namely Colonel Karl Friedl and Lieutenant Colonel Michael Russo of the U.S. Army Medical Research and Materiel Command in Fort Detrick, MD. We also must acknowledge the invaluable contributions of Major Brandon Doan of the U.S. Air Force

Research Laboratory who provided an unbelievable amount of behind-the-scenes support during every phase of this effort from protocol approval to report completion. Last but not least, we are deeply grateful to the F-117 pilots who selflessly dedicated their time to serve as research participants. Their enthusiasm, skill, and professionalism ensured that the data collected were of the highest possible quality, and they made the entire study possible.

REFERENCES

Akerstedt, T. 1995. Work hours, sleepiness and the underlying mechanisms. Journal of sleep research, 4(suppl. 2):15-22.

Balkin, T., Thorne, D., Sing, H., Thomas, M., Redmond, D., Wesensten, N., Williams, J., Hall, S., and Belenkey, G. 2000. Effects of sleep schedules on commercial motor vehicle driver performance. Report no. DOT-MC-00-133. Washington, DC: Department of Transportation Federal Motor Carrier Safety Administration.

Belenky, G., Penetar, D. M., Thorne, D., Popp, K., Leu, J., Thomas, M., Sing, H., Balkin, T., Wesensten, N., and Redmond, D. 1994. The effects of sleep deprivation on performance during continuous combat operations. In Food Components to Enhance Performance. Washington, DC: National Academy Press; 127-35.

Caldwell, J. A., Caldwell, J. L., and Crowley, J. S. 1996. Sustaining helicopter pilot alertness with Dexedrine during sustained operations. Proceedings of the Advisory Group for Aerospace Research and Development, Aerospace Medical Symposium on Neurological Limitations of Aircraft Operations: Human Performance Implications, CP -579. Neuilly Sur Seine, France: North Atlantic Treaty Organization.

Caldwell, J. A., Caldwell, J. L., Smythe, N. K., and Hall, K. K. 2000a. A double-blind, placebo-controlled investigation of the efficacy of modafinil for sustaining the alertness and performance of pilots: a helicopter simulator study. Psychopharmacology, 150:272-282.

Caldwell, J. A., Hall, K. K., and Erickson, B. S. 2002. EEG data collected from helicopter pilots in flight are sufficiently sensitive to detect increased fatigue from sleep deprivation. International journal of aviation psychology, 12(1):19-32.

Caldwell, J.A., Hall, K.K., Erickson, B., Labrada, E. and Norman, D. 2001. The sensitivity of real-time telemetered electroencephalographic data and actual in-flight performance measurement to fatigue in a sleep deprivation paradigm. USAARL Technical Report No. 2001-02. Fort Rucker, AL: U. S. Army Aeromedical Research Laboratory.

Caldwell, J. A., Kelly, F., Roberts, K. , Jones, H., Lewis, J., Woodrum, L., Dillard, R. and Johnson, P. 1997. A comparison of EEG and evoked response data collected in a UH-1 helicopter to data collected in a standard laboratory environment. USAARL Technical Report No. 97-30. Fort Rucker, AL: U. S. Army Aeromedical Research Laboratory.

Caldwell, J. A., Lewis, J. A., Darling, S. R., Dillard, R. M., and Johnson, P. P. 1994a. Collection of real-time, multichannel EEG data from helicopter pilots in flight: A feasibility study. USAARL Technical Report No. 94-26. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory.

Caldwell, J. A., Prazinko, B. F., and Caldwell, J. L. 2003. Body posture affects electroencephalographic activity and psychomotor vigilance task performance in sleep-deprived subjects. Clinical neurophysiology, 114:23-31.

Caldwell, J. A., Smythe, N. K., LeDuc, P. A., and Caldwell, J. L. 2000b. Efficacy of dextroamphetamine for the maintenance of pilot performance during 64 hours of sustained wakefulness. Aviation, space, and environmental medicine, 71(1):7-18.

Caldwell, J. A., Wilson, G. F., Muzaffer, C., Gaillard, A. W. K., Gundel, A., Lagarde, D., Makeig, S., Myhre, G., and Wright, N. A. 1994b. Psychophysiological assessment methods. NATO AGARD Advisory Report No. 324. Neuilly Sur Seine, France: North Atlantic Treaty Organization.

Comperatore, C. A., Caldwell, J. A., Stephens, R. L., Mattingly, A., Chiaramonte, J., and Trast, S. T. 1993. The use of electrophysiological and cognitive variables in the assessment of degradation during periods of sustained wakefulness. USAARL Technical Report No. 93-5. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory.

Comstock, J. R., and Arnegard, R. J. 1992. The multi-attribute task battery for human operator workload and strategic behavior research. Hampton, NY: Langley Research Center, National Aeronautics and Space Administration. NASA Technical Memorandum No. 104174.

Dawson, D., and Reid, K. 1997. Fatigue, alcohol and performance impairment. Nature, 388:235.

Dinges, D. F. 1990. The nature of subtle fatigue effects in long-haul crews. Proceedings of the Flight Safety Foundation 43rd International Air Safety Seminar; Rome, Italy. Arlington, VA: Flight Safety Foundation.

Dinges, D. F. 1989. The nature of sleepiness: Causes, contexts and consequences. In Stunkard, A., Baum, A. (Eds) Perspectives in behavioral medicine: Eating, sleeping, and sex. Hillsdale: Lawrence Erlbaum.

Dinges, D. F., Mallis, M. M., Maislin, G., Powell, J.W. 1998. Evaluation of Techniques for Ocular Measurement as an Index of Fatigue and as the Basis for Alertness Management, Report no. DOT HS 808 762, April, 1998:42-47.

Elliot, S. 2001. Chief of staff shares views on Global Strike Task Force. Air Force news archive. www.af.mil/news/Oct2001/n20011029_1543.shtml.

Horne, J. A. 1978. A review of the biological effects of total sleep deprivation in man. Biological psychology, (7):55-102.

Horne, J. A., and Reyner, L. A. 1995. Selep related vehicle accidents. British Medical Journal, 310:565-567.

Howitt, J. S., Hay, A. E., Shergold, G. R., and Ferres, H.M. 1978. Workload and fatigue in-flight EEG changes. Aviation, space, and environmental medicine, 49:1197-1202.

Kirsch, A. D. 1996. Report on the statistical methods employed by the U.S. FAA in its cost benefit analysis of the proposed "Flight Crewmember Duty Period Limitations, Flight Time Limitations and Rest Requirements", Docket No. 28081. Comments of the Air Transport Association of America to FAA notice 95-18, FAA Docket No. 28081. Appendix D, pp. 1-36.

Klein, D. E., Bruner, H., and Holtman, H. 1970. Circadian rhythm of pilot's efficiency, and effects of multiple time zone travel. Aerospace medicine, 41:125-132.

Krueger, G. P. 1989. Sustaining military performance in continuous operations: Combatant fatigue, rest and sleep needs. Handbook of military psychology New York: John Wiley and Sons, 255-277.

L-3 Communications. 1993. Operation and maintenance instructions, F-117 Weapon System Trainer (WST), Lockheed Martin Skunk Works Technical Manual no. TO43D3-4-117-1-1. Binghamton, NY: L-3 Communications.

LaFontaine, E., and Medvedeff, M. 1966. Radiotelemetric recordings of the electroencephalograms of civil aviation pilots during flight. Clinical aviation and aerospace medicine, October:1060-1065.

LeDuc, P. A., Caldwell, J. A., and Ruyak, P. S. 2000. The effects of exercise as a countermeasure for fatigue in sleep-deprived aviators. Military psychology, 12(4):249-266.

Lorenzo, I., Ramos, C. A., Guevara, M. A., Corsi-Cabrera, M. 1995. Effect of total sleep deprivation on reaction time and waking EEG activity in man. Sleep, 18(5):346-354.

Maulsby, R. L. 1966. Electroencephalogram during orbital flight. Aerospace medicine, October: 1022-1026.

McNair, D. M., Lorr, M., and Droppleman, L. F. 1981. Manual for the profile of mood states. San Diego: Educational and Industrial Testing Service.

Mesalum, M. 1985. Principles of behavioral neurology. Philadelphia: F.A. Davis Company.

Moore-Ede, M. 1993. Aviation safety and pilot error, In Twenty-four hour society, pp. 81-95. Reading, MA: Addison-Wesley Publishing Co.

Neri, D. F., Oyung, R. L., Colletti, L. M., Mallis, M. M., Tam, P. Y., and Dinges, D. F. 2002. Controlled breaks as a fatigue countermeasure on the flight deck. Aviation, space, and environmental medicine, 73:654-664.

Penetar, D., McCann, U., Thorne, D., Kamimori, G., Galinski, C., Sing, H., Thomas, M., and Belenky, G. 1993. Caffeine reversal of sleep deprivation effects on alertness and mood. Pscyhopharmacology, 112:359-365.

Perry, I. C. 1974. Helicopter aircrew fatigue. AGARD Advisory Report No. 69. Neuilly sur Seine, France: Advisory Group for Aerospace Research and Development.

Pigeau, R. A., Heselegrave, R. J., and Angus, R. G. 1987. Psychophysiological measures of drowsiness as estimators of mental fatigue and performance degradation during sleep deprivation. In Electric and magnetic activity of the central nervous system: Research and clinical applications in aerospace medicine. AGARD CP-432, 21-1/21-16. Neuilly sur Seine, France: Advisory Group for Aerospace Research and Development.

PMI, Inc. (1999). FIT Workplace Safety Screener: Theory and validation. Rockville, MD: PMI, Inc.

Ramsey, C. S, and McGlohn, S. E. 1997. Zolpidem as a fatigue countermeasure. Aviation, space, and environmental medicine, 68(10):926-931.

Rechtschaffen, A., and Kales, A. (Eds.) 1968. A manual of standardized terminology, techniques, and scoring system of human subjects. NIH publication no. 204. Washington, DC: U.S. Government Printing Office.

Reeves, D., Kane, R., Winter, K., Ransford, K., and Pancella, R. 1993. Automated Neuropsychological Assessment Metrics (ANAM): Test Administration Guide, Version 1.0. St. Louis, Missouri: Institute of Mental Health.

Rosekind, M. R., Gregory, K. B., Co, E. L., Miller, D. L., and Dinges, D. F. 2000. Crew factors in flight operations XII: A survey of sleep quantity and quality in on-board crew rest facilities. NASA Technical Memorandum no. 2000-209611. Moffett Field, CA: National Aeronautics and Space Administration, Ames Research Center.

Rosekind, M. R., Graeber, R. C., Dinges, D. F., Connell, L. J., Rountree, M. S., Spinweber, C. L., Gillen, K. A. 1994. Crew factors in flight operations IX: Effects of planned cockpit rest on crew performance and alertness in long-haul operations. NASA Technical

Memorandum no. 108839. Moffet Field, CA: National Aeronautics and Space Administration, Ames Research Center.

Russo, M., Thomas., Sing, H., Thorne, D., Balkin, T., Wesensten, N., Redmond, D., Welsh, A., Rowland, L., Johnson, D., Cephus, R., Hall, S., Belenky, G. & Krichmar, J. 1999. Sleep deprivation related changes correlate with simulated motor vehicle crashes. pp. 119-127. In R. Carroll (Ed.) Ocular measures of driver alertness: Technical conference proceedings, FHWA Technical Report No. FHWA-MC-99-136. Washington, DC: Federal Highway Administration, Office of Motor Carrier and Highway Safety.

Russo, M., Thomas, M., Thorne, D., Sing, H., Redmond, D., Rowland, L., Johnson, D., Hall, S., Krichmar, J., and Balkin, T. 2003. Oculomotor impairment during chronic partial sleep deprivation. Clinical neurophysiology, 114:723-726.

Sterman, M. B., Schummer, G. J., Dushenko, T. W., and Smith, J. C. 1987. Electroencephalographic correlates of pilot performance: Simulation and in-flight studies. In Electrical and magnetic activity of the central nervous system: Research and clinical applications in aerospace medicine, AGARD CP No. 432, 31-1/31-16. Neuilly Sur Seine, France: NATO.

Stern, J. A. 1980. Aspects of visual search activity related to attentional processes and skill development, Final Report, Contract no. F49620-79-C0089. Washington, DC: Air Force Office of Scientific Research.

Stern, A. 1999. Ocular based measures of driver alertness. In Ocular measures of driver alertness: Technical conference proceedings. pp. 4- 9. Herndon, VA: Federal Highway Administration Office of Motor Carrier and Highway Safety, National Highway Traffic Safety Administration Office of Vehicle Safety Research.

Wierwille, W. 1999. Historical perspective on slow eyelid closure: Whence PERCLOS? In Ocular measures of driver alertness: Technical conference proceedings. pp. 31-51. Herndon, VA: Federal Highway Administration Office of Motor Carrier and Highway Safety, National Highway Traffic Safety Administration Office of Vehicle Safety Research.

Wildzunas, R. M. 1999. Notes on the calculation of power in determining sample size and the construction of digram-balanced designs. USAARL Technical Report No. 99-14. Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory.